IEMTRONICS 2020

International IOT, Electronics and Mechatronics Conference

Vancouver, Canada

Conference Proceedings

9th-12th September 2020

SPONSORS:

IEM
अश्नावान लभते ज्ञानम्
INSTITUTE OF ENGINEERING & MANAGEMENT
Good Education, Good Jobs

IEEE VANCOUVER SECTION

SMART

UEM
अश्नावान लभते ज्ञानम्
UNIVERSITY OF ENGINEERING & MANAGEMENT
Good Education, Good Jobs

Editors:

Satyajit Chakrabarti, Rajashree Paul, Bob Gill, Malay Gangopadhyay, Sanghamitra Poddar

About the Conference

IEMTRONICS 2020

Continuing with the outstanding success of IEEE IEMCON, IEEE CCWC, IEEE UEMCON, IEMANTENNA we are proud to present IEMTRONICS 2020 (International IOT, Electronics and Mechatronics Conference) which will be held during **9th -12th September, 2020 at Vancouver, Canada, in online mode**. Keeping in mind the pandemic situation prevalent globally due to Covid 19 and following the legacy of organizing highly successful conferences, we have planned for the online conference. The conference aims to bring together scholars from different backgrounds to emphasize dissemination of ongoing research broadly in the fields of IOT, Electronics and Mechatronics. Research papers were invited describing original works in above mentioned fields and related technologies. The conference will include a peer-reviewed program of technical sessions, special sessions, tutorials and demonstration sessions.

All accepted papers which will be presented during the parallel sessions of the Conference will be published in IEEE Xplore Digital Library (Scopus, DBLP, Ei Compendex, Web of Science and Google Scholar indexed).

This conference will also promote an intense dialogue between academia and industry to bridge the gap between academic research, industry initiatives, and governmental policies. This is fostered through panel discussions, keynotes, invited talks and industry exhibits where academia is exposed to state-of-practice and results from trials and interoperability experiments. The industry in turn benefits by exposure to leading-edge research in networking as well as the opportunity to communicate with academic researchers regarding practical problems that require further research.

Our Reviewers

IEMTRONICS 2020 followed a rigorous triple-blind review process in order to identify suitable papers for both presentation and publication. This process helped the organizers to shortlist good quality papers from diverse regional areas and across various domains. This edited book also incorporates four invited papers from experts across the globe. The congress received more than three hundred full papers for review and approximately twenty five percent were selected for full paper submission. In the end, eighty papers, including invited papers were found acceptable for presentation and congress proceedings.

Such a detailed review process was possible due to the excellent and enthusiastic support extended by the strong technical review team of IEMTRONICS 2020. For every stage of submission, IEMTRONICS had a specific template review procedure to analyze the submissions and provide suitable comments for the authors to incorporate. The review team which formed the technical backbone for the selection of submissions for the edited book and the conference presentation was supervised by::

- Mr. Qasem Abu Al-Haija
Tennessee State University

- Dr. Mohd Ashraf Ahmad
Universiti Malaysia Pahang

- Dr. Ritesh Ajoodha
The University of the Witwatersrand, Johannesburg

- Dr. Md Ali
Rider University

- Dr. Mohammad AlShabi
University of Sharjah

- Prof. Tarachand Amgoth
Indian Institute of Technology (ISM) Dhanbad

- Mr. Mohammad Anees
Xilinx

- Mrs. Priyalakshmi Balasundaram
SRM Institute of Science and Technology

- Dr. Sankhyabrata Bandyopadhyay
Southern University of Science and Technology

- Prof. Aleksandr Belov
National Research University Higher School of Economics

- Dr. Ankur Bhattacharjee
Birla Institute of Technology and Science Pilani, Hyderabad Campus

- Prof. Rajat Subhra Chakraborty
Indian Institute of Technology

- Dr. Pratik Chattopadhyay
Indian Institute of Technology (BHU), Varanasi

- Mr. Sangay Chedup
Jigme Namgyel Engineering College

- Dr. Khalid Darabkh
The University of Jordan

- Mr. Arighna Deb
KIIT University

- Dr. Arindam Deb
KIIT

- Dr. Rajiv Dey
SoET, BML Munjal University

- Mr. Sukomal Dey
Indian Istitute of Technology, Delhi

- Dr. Ozan Dogan
Delft

- Dr. Ke-Lin Du
Concordia University

- Dr. Tarek El Salti
Sheridan College

- Dr. Heba Elgazzar
Morehead State University

- Ms. Eiman Elghanam
American University of Sharjah

- Mr. Vipkas Al Hadid Firdaus
State Polytechnic of Malang

- Dr. Stephen Andrew Gadsden
University of Guelph

- Dr. Vilas Gaidhane
Birla Institute of Technology and Science Pilani, Dubai Campus

- Mr. Nilesh Goel
Birla Institute of Technology and Science Pilani, Dubai Campus, Dubai UAE

- Dr. Arnab Hazra
BITS Pilani Rajasthan

- Dr. Ranjay Hazra
Nit Silchar

- Ms. Mousam Hossain
University of Central Florida

- Dr. Rasha Kashef
Ryerson University

- Mr. Hiroshige Kikura
Tokyo Institute of Technology

- Prof. Jong-Hoon Kim
Kent State University

- Mr. Ashish Kulkarni
DTU Delhi

- Dr. Kean Boon Lee
The University of Sheffield

- Dr. Moises Levy
West Texas A&M University

- Dr. Sudipta Maity
National Institute of Technology (NIT), Rourkela

- Dr. Pravir Malik
Deep Order Technologies

- Dr. Morteza Modarresi Asem
Tehran Medical Sciences University

- Mr. Anandarup Mukherjee
University of Cambridge

- Dr. Zia Nadir
Sultan Qaboos University

- Dr. Ibrokhimali Normatov
National University of Uzbekistan named after Mirzo Ulugbek

- Mr. Kumar Rahul
XILINX

- Dr. Ashiq Sakib
Florida Polytechnic University

- Dr. Angsuman Sarkar
Kalyani Government Engiuneering College

- Mr. Ingmar Schwarz
TU Dortmund

- Dr. Mrinal Sen
Indian Institute of Technology(ISM), Dhanbad

- Dr. Rohit Singh
Carnegie Mellon University

- Dr. Kanika Sood
California State University, Fullerton

- Mr. Dodit Suprianto
Politeknik Negeri Malang

- Mr. Sourabh Swarnkar
Xilinx

- Dr. Shahab Tayeb
California State University, Fresno

- Dr. Girma Tewolde
Kettering University

- Dr. Mustafa Tunay
Istanbul Gelisim University

- Prof. Damian Valles
Texas State University

- Dr. Kristiina Valter Mai
Ryerson University

- Dr. Aparna Varde
Montclair State University

- Mr. Santosh Yachareni
Xilinx Semiconductor

- Dr. Hasan Yasar
Carnegie Mellon University

Sponsors

Society for Makers, Artists, Researchers and Technologists, Canada

IEEE VANCOUVER SECTION

Institute of Engineering & Management, Kolkata

University of Engineering & Management, Kolkata

University of Engineering & Management, Jaipur

COPYRIGHT

2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS).

IEEE Catalog Number: **CFP20Y72-ART**

ISBN: 978-1-7281-9615-2

Organizing Committee:

General Chair:

- **Rajashree Paul**

  University of Engineering & Management, Kolkata, India

Technical Co-Chair:

- **Bob Gill**

  British Columbia Institute of Technology, Burnaby, Canada

- **Malay Gangopadhyay**

  Institute of Engineering & Management, Kolkata, India

Publicity Chair:

- **Sanghamitra Poddar**

  Institute of Engineering & Management, Kolkata, India

Advisory Committee

Name	University
Dr. Ashutosh Datta	Johns Hopkins University, USA
Dr. Yang Hao	Queen Mary University, London
Dr. Vien Van	University of Alberta, Canada
Dr. Omar Ramahi	University of Waterloo, Canada
Dr. Yahia Antar	Royal Military College, Canada
Dr. Zhizhang (David) Chen	Dalhousie University, Canada
Dr. Detlef Streitferdt	Technische Universitat Ilmenau, Germany

Technical Committee

Name	University
Dr. Nabeeh Kandalaft	Grand Valley State University, USA
Dr. Alex "Sandy" Antunes	Capitol Technology University, USA
Dr. Izzat Alsmadi	Texas A&M, San Antonio, USA
Dr. Lo'ai Tawalbeh	Texas A&M University-San Antonio, USA
Dr. Pratik Chattopadhyay	Indian Institute of Technology (BHU), Varanasi
Dr. Doina Bein	California State University, Fullerton, USA
Dr. Hasan Yasar	Carnegie Mellon University, USA
Dr. Moises Levy	Florida Atlantic University, USA
Dr. Christian Trefftz	Grand Valley State University, USA
Dr. Mrinal Sen	Indian Institute of Technology(ISM), Dhanbad
Dr. Petros Spachos	University of Guelph, Canada
Dr. Kanika Sood	California State University, Fullerton
Dr. Ke-Lin Du	Concordia University, Canada
Dr. Wenlin Han	California State University, Fullerton
Dr. Ashiq Adnan Sakib	Florida polytechnic University, USA
Dr. Morteza Modarresi Asem	Islamic Azad University, Iran
Dr. Md. Liakat Ali	Rider University, USA
Dr. Tarek El Salti	Sheridan College, Canada
Dr. Sukomal Dey	Indian Institute of Technology, Palakkad
Dr. Maysam Chamanzar	Carnegie Mellon University, USA
Dr. Shahab Tayeb	California State University, Fresno
Dr. Kean Boon Lee	Sheffield University, UK

Tracks:

Track 1: IoT & Data Science

- **IoT and blockchain**
- **IoT and big data**
- **Next-generation infrastructure for IoT**
- **Cloud computing and IoT**
- **Edge computing and IoT**
- **IoT platforms, tools, and applications**
- **IoT systems development methodologies**
- **IoT applications**

Track 2: Electronics

- **Photonics**
- **Nano science & quantum technology**
- **VLSI and microelectronic circuit embedded systems**
- **System on Chip (SoC) Design**
- **FPGA (Field Programmable Gate Array) design and applications**
- **Electronic instrumentations**
- **Sensors & systems**
- **NEMS & MEMS**
- **Integrated circuits & power electronics**
- **Electronic power converters and inverters**
- **Electric vehicle technologies**
- **Control theory, optimization and applications**
- **Robotics and autonomous systems**
- **Intelligent,optimal,robust,adaptive control**
- **Linear and nonlinear control systems**
- **Complex adaptive systems**
- **Industrial automation and control systems technology**
- **Modern electronic devices**
- **Biomedical devices & imaging**
- **Energy harvesting & conversions**
- **Energy efficient hardware systems**

Track 3: Mechatronics

- **Sensing and control systems**
- **Mechatronics systems**
- **Mechanical systems**
- **Artificial intelligence**
- **Applications of robotics**

Track 4: Information Technology

- **Business intelligence and applications**
- **Computer network**
- **Evolutionary computation and algorithms**
- **Intelligent information processing**
- **Information system integration and decision support**
- **Image processing and multimedia technology**
- **Signal detection and processing**
- **Technique and application of database**
- **Software engineering**
- **Mobile computing**
- **Distributed systems**
- **Artificial intelligence**
- **Visualization and computer graphic**
- **Natural language processing**
- **Machine learning**
- **Internet of things, data mining**
- **Data science**
- **Cloud computing in e-commerce scenarios**
- **E-business systems integration and standardization, e-government**
- **Electronic business model and method**
- **E-commerce risk management**
- **Recommender system**
- **Semantic web service architecture for e-commerce**
- **Service oriented e-commerce & business processes**
- **Data analytics and big data**
- **Software defined networking**
- **Secured distributed systems**

Track 5: Mobile Communication

- **Ad hoc networks**
- **Body and personal area networks**
- **Cloud and virtual networks**
- **Cognitive radio networks**
- **Cyber security**
- **Cooperative communications**
- **Delay tolerant networks**
- **Future wireless internet**
- **Local dependent networks**
- **Location management**
- **Mobile and wireless IP, mobile computing**
- **Multi-hop networks**
- **Network architectures**
- **Network Security, information security, encoding technology**
- **Routing, QoS and scheduling**
- **Satellite communications**
- **Self-organising networks**
- **Telecommunication Systems**
- **Vehicular networks**
- **Wireless multicasting, wireless sensor networks**

Dr. Vincent Chan

Massachusetts Institute of Technology

Bio : Vincent W. S. Chan, the Joan and Irwin Jacobs Professor of EECS, MIT, received his BS(71), MS(71), EE(72), and Ph.D.(74) degrees in EE all from MIT. From 1974 to 1977, he was an assistant professor, EE, at Cornell University. He joined MIT Lincoln Laboratory in 1977 and had been Division Head of the Communications and Information Technology Division until becoming the Director of the Laboratory for Information and Decision Systems (1999–2007). He is currently a member of the Claude E. Shannon Communication and Network Group at the Research Laboratory of Electronics of MIT.

In July 1983, he initiated the Laser Intersatellite Transmission Experiment Program and in 1997, the follow-on GeoLITE Program. In 1989, he formed the All-Optical-Network Consortium among MIT, AT&T and DEC. He also formed and served as PI the Next Generation Internet Consortium, ONRAMP among AT&T, Cabletron, MIT, Nortel and JDS, and a Satellite Networking Research Consortium formed between MIT, Motorola, Teledesic and Globalstar. He has founded in 2009 and is serving as the Editor-in-Chief of a new IEEE/OSA Journal: Journal of Optical Communications and Networking. He is currently a Member of the Corporation of Draper Laboratory, the Technical Advisory Board of Mercury Computer and on the Board of Governors of the IEEE Communication Society. He is also an elected member of Eta-Kappa-Nu, Tau-Beta-Pi and Sigma-Xi, the Fellow of the IEEE and the Optical Society of America.

Throughout his career, Professor Chan has spent his research focus on communication and networks, particularly on free space and fiber optical communication and networks and satellite communications. His work has led the way to a successful laser communication demonstration in space and early deployment of WDM optical networks. His recent research emphasis is on heterogeneous (satcom, wireless and fiber) network architectures with stringent performance demands.

Dr. William Lehr

Massachusetts Institute of Technology

Bio : Dr. William Lehr is a telecommunications and Internet industry economist and consultant with over twenty-five years of experience. He regularly advises senior industry executives and policymakers in the U.S. and abroad on the market, industry, and policy implications of events relevant to the

Internet ecosystem. He is a research scientist in the Computer Science and Artificial Intelligence Laboratory (CSAIL) at the Massachusetts Institute of Technology, currently engaged several multidisciplinary research projects within the Advanced Networking Architecture Group in CSAIL (ANA). Dr. Lehr's research focuses on the economics and regulatory policy of the Internet infrastructure industries. He teaches courses on the economics, business strategy, and public policy issues facing telecommunications, Internet, and eCommerce companies, and is a frequent speaker at international industry and academic conferences. He is the author of numerous publications on such topics as the measurement of economic impacts of Information technologies, the economics of technical standard setting, the pricing of Internet services, and the implications of commercializing novel Internet and wireless technologies for industry structure and regulatory policy.

In addition to his academic research, Dr. Lehr provides litigation, economic, and business strategy consulting services for firms in the information technology industries in the U.S. and abroad. Dr. Lehr has advised information technology companies on strategic marketing, pricing, financial planning, and competitive strategy; and government agencies in the United States and abroad on telecommunications and Internet policy matters. Dr. Lehr has prepared expert witness testimony for both private litigation and for regulatory proceedings before the FCC, before numerous state commissions and for numerous regulatory agencies abroad.

Dr. Lehr holds a PhD in Economics from Stanford (1992), an MBA from the Wharton Graduate School (1985), and MSE (1984), BS (1979) and BA (1979) degrees from the University of Pennsylvania.

Dr. Chuck Easttom

University of Dallas, USA & Georgetown University, USA

Research Interest: Cryptography, Cyber Warfare, Engineering Processes and Digital Forensics

Bio : Dr. **Chuck Easttom** is the author of 29 books, including several on computer security, forensics, and cryptography. He has also authored scientific papers on digital forensics, cyber warfare, cryptography, and applied mathematics. He is an inventor with 20 computer science patents. He holds a Doctor of Science (D.Sc.) in cyber security (dissertation topic: a study of lattice-based algorithms for post quantum cryptography) and a Doctor of Philosophy (Ph.D.) in Technology focused on bioengineering and nanotechnology. Dissertation topic is "The effects of nonlinear dynamics on nanotechnology and bioengineering" as well as three master's degrees (one in applied computer science and one in systems engineering). He also has 55 industry certifications (CISSP, CEH, etc.) He has been active in the IEEE Systems and Software Engineering Standards Committee in the 2675 DevOps group and the P2731 Brain Computer Interface standards group, as well as a reviewer for IEEE Open Access. He is a Distinguished Speaker of the ACM and

senior member of the ACM as well as a Distinguished Visitor of the IEEE and an IEEE Senior Member. Dr. Easttom is an adjunct lecturer for Georgetown University teaching graduate cyber security courses in their Master of Professional Studies in Applied Intelligence program and an adjunct professor for the University of Dallas teaching digital forensics.

Dr. Phillip Bradford

University of-Connecticut-Stamford, USA

Research Interest: Algorithms, Computer security, IoT, Blockchain, Artificial intelligence

Bio : Dr. **Bradford** is a computer scientist with extensive experience in academia and industry. He has great reverence for quality research. He believes engineering and applied science is the main way to grow the economy by improving our living standards. Phil was a post-doctoral fellow at the Max-Planck-Institut für Informatik, he earned his PhD at Indiana University, an MS form the University of Kansas, and a BA from Rutgers University. He was on the faculty at Rutgers Business School and the University of Alabama School of Engineering. He has worked for General Electric, BlackRock, Reuters Analytics, founded a firm and both occasionally consults and works with a number of early stage firms. Currently, Phil is the director of the computer science program at the University of Connecticut, Stamford.

Dr. Ronald F. DeMara

University of Central Florida

Bio : Ronald F. DeMara is Pegasus Professor in the Department of Electrical and Computer Engineering, joint faculty of Computer Science, and the Digital Learning Faculty Fellow at the University of Central Florida, where he has been a full-time faculty member since 1993. His interests are in computer architecture, post-CMOS devices, and reconfigurable fabrics. He has applied these to autonomous, embedded, and intelligent/neuromorphic systems, on which he has completed approximately 300 articles, 50

funded projects as PI or Co-PI, and 50 graduates as Ph.D. dissertation and/or M.S. thesis advisor. He was previously an Associate Engineer at IBM and a Visiting Research Scientist at NASA Ames, in total for four years, and is a registered Professional Engineer since 1992. He has served ten terms as a Topical Editor or Associate Editor including IEEE Transactions on Computers, Transactions on Emerging Topics in Computing, Transactions on VLSI, IEEE Spectrum, and Technical Program Committees of various IEEE conferences. He has been Keynote Speaker of IEEE RAW and IEEE ReConFig conferences, and Guest Editor of IEEE Transactions on Computers 2017 Special Section on Innovation in Reconfigurable Fabrics and 2019 Special Section on Non-Volatile Memories. He is currently an Associate Editor of IEEE Transactions on Emerging Topics in Computing and an IEEE Spectrum Editorial Advisory Board Member. He received the Joseph M. Biedenbach Outstanding Engineering Educator Award from IEEE in 2008.

Dr. Scott C. Smith

Texas A&M University-Kingsville

Bio : Dr. Scott C. Smith received B.S. degrees in Electrical Engineering and Computer Engineering, and an M.S. degree in Electrical Engineering, from the University of Missouri – Columbia, in 1996 and 1998, respectively, and a Ph.D. degree in Computer Engineering from the University of Central Florida, Orlando, in 2001. He started as an Assistant Professor at University of Missouri – Rolla (now called Missouri University of Science & Technology), where he was promoted to Associate Professor with tenure, before moving to University of Arkansas (UA) as a tenured Associate Professor. After 6 years at UA, Dr. Smith moved to North Dakota State University as a Full Professor to become ECE Department Chair; and he is now Professor and Chair of the Department of Electrical Engineering and Computer Science at Texas A&M University – Kingsville. Dr. Smith has received more than $4.8 million in research funding from various government agencies and private industry, and has co-founded 3 technology based startup companies. He has authored over 100 refereed journal and conference publications, plus 8 U.S. patents, 1 co-authored and another co-edited book, and 4 book chapters. Dr. Smith graduated 12 Ph.D. and 14 MS students, and supervised over 100 undergraduate research students. His research interests include Asynchronous Logic, NULL Convention Logic (NCL), Computer Architecture, Embedded Systems, Digital Logic, FPGAs, CAD Tools for Digital Design, Computer Arithmetic, VHDL, VLSI, Engineering Education, Secure/Trustable Hardware, Wireless Sensor Networks, Robotics, and Cyber Physical Systems. Dr. Smith is a Senior Member of IEEE, and a Member of the National Academy of Inventors, Sigma Xi, IEEE-Eta Kappa Nu, Tau Beta Pi, and ASEE.

Dr. Gideon Samid

CTO, BitMint
Case Western Reserve University

Bio : Born in Jerusalem, growing up together with his young country, Israel, Gideon was inspired by the immensity of a 2000 years dream becoming a reality, and by a desert made to bloom and shine, while facing hostilities from every direction. And ever since, he was thinking horizons, possibilities, projections, understanding. His high school motif was: "Be Humble". It exercised military style uniform and discipline. Gideon served as an infantry intelligence officer, and studied engineering at the Technion — Israel institute of Technology. He worked for the government, and in the oil and energy industry. After the trauma of the Yom Kippur war Gideon ventured abroad. He was selected to represent Israel on a NASA project, which evolved into a long standing engineering and innovation career threaded through Exxon, SAC, Presearch, The Pentagon, and D&G Science — Innovation Productivity Corporation. Gideon has found his niche: developing the general theory of innovation: mathematical and logical tools to appraise, measure, and minimize the effort required to achieve an innovative goal. Gideon returned to the Technion to complete his PhD dissertation, and perfect these tools, and then he applied them on a variety of applications ranging from innovative industrial contraptions, through a powerful inference engine (BiPSA), cryptographic data representation, and lately: design and development of digitized money through advanced chemistry and cryptography (patents awarded).

Gideon Samid is a natural born teacher. He enjoys his face to face students at the University of Maryland, his online teaching with the American Chemical Society, his work in the department of computer science at Case Western Reserve University, and his innovative hybrid teaching developed for Class Express, LLC.

Gideon work on appraising innovative efforts through quantifying the measure of 'relevant unknown' dove-tails with his parallel philosophical pursuit of the notion of Unbound Ignorance — the idea that most of what we need to know, we don't, and most of what we need to forget we won't.
What we miss is not a plank in an otherwise well established platform, but it is a platform we had not yet extended our imagination to comprehend. This realization led to a behavioral and life-decisions guide: the idea of try again, since you are not 'there' yet. Gideon employs all his writing and artistic talent to express this very point. When asked to summarize Gideon says: Act Rationally, Live Poetically.

Books: in "Computer Organized Cost Engineering" Gideon presents the methods to appraise missing knowledge via the credibility of the estimate of cost to complete, and time to finish. In "The Innovation Turing Machine" one finds the essentials of the general theory of innovation. In "The Unending Cyber War" Gideon unveils a cyber security approach based on appraising the missing knowledge from the point of view of the attacker, and how to keep this knowledge gap

as large as possible — reducing these principles to a fabric of practical procedures and protocols. In “Fair Moon” the reader will find some forty-six story lessons describing the effort to acknowledge life uncertainties. In “Unbound Ignorance” — the grand admission of our ignorance of life and reality is laid out. And in his latest book “Tethered Money” Gideon presents the grand vision for the future of money — digitized and tethered to its intended purpose, resisting waste, fraud and abuse. What a vision! It keeps Gideon busy as Chief Technology Officer for BitMint — the company that developed this bit-wise money technology.

Dr. Ali Kanso

Senior Cloud Software Engineer, IBM

Bio : Ali Kanso is a senior software engineer in cloud technologies at IBM T.J. Watson research center, NY. He has developed HA (high availability) solutions for cloud-hosted applications. His main area of expertise is distributed systems including cloud management systems (such as OpenStack), and container management frameworks (such as Kubernetes and Docker Swarm), as well as virtual synchrony algorithms and their implementations. His research focuses on developing HA solutions to monitor, isolate, repair and analyze the systems’ behavior under faulty and disastrous conditions. Prior to IBM, he was a senior researcher at Ericsson, where he also served at the young advisory board committee advising on Ericsson research strategy. Ali Kanso also an adjunct research professor at the University of Western Ontario (WesternU).

Content:

Ring-Shaped Force Platform for Developing Instrumented Throwing and Shifting Forces Due to Rotation Analysis

Maher J. Ameen
Department of Physical Education
University of Kufa
Al-Najaf, Iraq
Maher.shlash5@gmail.com

Abdulkadhim A. Hassan
Department of Electrical Engineering
University of Kufa
Al-Najaf, Iraq
abdulkadhim.shlash@uokufa.edu.iq

Ahmed M. Sadeq Al-Salakh
Department of Civil Engineering
University of Kufa
Al-Najaf, Iraq
ahmedm.alsalakh@uokufa.edu.iq

Janan H. Saadie
Department of Electrical Engineering
University of Kufa
Al-Najaf, Iraq
Jenan.saadie@uokufa.edu.iq

Ammar M. Alnajem
Department of Physical Education
University of Kufa
Al-Najaf, Iraq
Ammarm.alnajem@uokufa.edu.iq

Abstract—**This paper presents further development and utilization of force platforms to evaluate and analysis of forces obtained due to rotation movements of sport activities using Labview software. The platform was designed with ring-shaped to provide and simulate the required space of rotation for testing these activities. The proposed force platform was fabricated with 2.5 m diameter and calibrated with standard force platforms to provide the required measurement sensitivity and accuracy. The data acquisition and Labview software were developed to perform different mathematical calculations. The obtained experimental tests results for different throwing sport activities showed good convergence to analyse athletes performances. Although the proposed force platform was initially employed to measure and analyse force due to rotation, it can be employed to measure other parameters related to athletes performances including torque, acceleration, displacement and moment based on computer vision tracking system. In comparison with the existing platforms, the fabricated platform has features of collecting a large variety of data with high quality for analysis and decision-making of athletes performances performing rotation movements.**

Keywords— Circular force platform, Throwing analysis, Force measurement, Calibration, Center of pressure

I. INTRODUCTION

All movements of biomechanics activities produce forces. To better understand and evaluate the mechanism of these movements, it is necessary to measure and examine the forces involved in such activities. Different techniques have been developed and used to achieve these purposes such as force platforms and other force measurement devices. Initially, the force platforms are designed to simultaneously measure, in three planes, external forces of ground reaction based on Newton's laws. Recently, force platforms are commonly employed for diagnosis and gait analysis purposes [1-3]. On the other hand, number of investigations have been conducted involving world-class sprinters [4-7]. Currently, the force platforms are used in the study of activities of jumping since they provide a direct measure of forces responsible for a jump to occur[8-9]. In all of the aforementioned applications, the geometrical shape of the used force plate is either square or rectangular shapes with small dimensions. For example, a square shape force platforms with dimensions of 0.6x0.6m are used in physics and sport [10]. Also, force plates with rectangular shape and dimensions of 0.914 x 0.46m are used to analyze the validity and reliability of a simple computation method to evaluate velocity, force , and output power during a jump countermovement [11-12].

The rotation events in the field and track such as hammer, discus and shot put throwing are technically very demanding. These activities involve performing complex movements with rotation in a limited space with high speed, which require a suitable force platform with specified size and circular shape. To the best of the author's knowledge, no appreciate research work was done to measure and analyze the forces and other parameters due to performing these activities. It worth's, however, to note that, despite the standard force platforms are useful, they would not be so powerful once force testing is going to be performed for these activities due to their limited sizes and shapes. In this paper, an instrumented force platform with ring-shaped geometry technique was proposed and fabricated to measure and analyze the forces due rotation movements of sport activities. The proposed platform was designed and manufactured with 2.5 m diameter to be used as mounted platform for laboratories applications purpose or as portable platform for field applications. It has features of measuring throwing forces and different parameters including torque, displacement, moment, and power of athletes. The force platform design was initiated and simulated using 3D-CAD software to define prototype geometry. The manufactured platform was first calibrated using two steps of calibration both static and dynamic calibrations. The calibrated force platform was subjected to different practical tests to evaluate its sensitivity and accuracy of measurement. The analysis of the measured parameters was performed using LabVIEW software as virtual instrument. Also, the system was enhanced with using multiple cameras with high speed and resolution for computer vision tracking system purpose to measure and to calculate different parameters.

II. FORCE PLATFORM DESIGN

A. Materils and Methods

The standard commercial force platform with rectangular is shown in Fig. 1 (a). For applications where forces due rotation movements, the standard force platforms are not suitable to be used for these purposes due to limitations of their sizes and shapes. Moreover, complex activities are involved with high speed during performing these activities. Hence, it becomes necessary to develop a

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

technique suitable to be used for measuring and analyzing different parameters related to such applications as shown in Fig. 1(b). The proposed platform is designed with a ring-shaped geometry to provide the required space with the specified diameter (D). It was first designed using CAD software to define its prototype geometry with circular area of 2.5 m diameter and 0.3 m height . This diameter of the platform represents the recommended standard value of the circle diameter required to perform different sport activities with rotational movement. The platform body structure was made from iron while the top floor was made from 10 mm-thick 5052F aluminium plate, weighting 6.30 kg, which was positioned on the structural foundation. The selection of this material can provide the platform with minimum mass and maximum rigidity to obtain a relatively high value of natural frequency and good linearity. The top surface of the platform was covered by 2.5 cm thickness of MDF wood to provide a smooth and non-skid surface without any disturbance that may be produced during rotation movements. Measurements were to be carried out using commercially available sensors (load cells) selected according to simplicity and criteria of robustness[13]. For the designed force platform, five load cells S_1, S_2, S_3, S_4, and S_5 (model SS300-2T, developed by SEWHACNM company, Korea) with total load capacity of 5000 N and sensitivity of 2 mV/V were used. The load cells were mounted around the top surface such that an angle of 72° is made between any two neighbour load cells as shown in Fig. 1(b) to provide the required accuracy in the measurements. Five screw drive were fixed at the lower surface of the platform to adjust the level of the plate on the ground and to adjust the load cell reading. This arrangement allows using the plate either to be used in the field or it can be fixed on concrete foundation in a laboratory. It is to be noted that the given specifications of the used load cells are put based on applying of vertical forces to the point of loading. The analogue electrical signals obtained from the sensors were collected via data logger data acquisition board. The acquired signals are then converted to digital data, which was fed to the laptop for further processing. The algorithms of measurement of force and other parameters were implemented using suitable dedicated LabVIEW lock routines for data storage and presentation developed for these purposes. The general block diagram of the measurement system is shown in Fig. 2 whiles Fig. 3 depict the measurement set up..

B. Calibration Procedure

Since force platforms only provide an electrical signal output as a voltage, calibration process is required to calculate the actual reading of load cells when they are subjected to different load conditions [14]. Two types of calibration were performed on the fabricated force platform, static calibration and dynamic calibration. For static case, the calibration process was performed in two steps. In the first step, the five load cells were first separately calibrated to access their linearity for a given rated range of applied loads. The second step of the calibration process was performed to access its operation when the whole measurement system is subjected to trial test loads (actual activities). The load cells static calibration was performed using calibrated loads with known values up to the rated load. The process of calibration was repeated and evaluated for each individual load cell with a load applied at the force platform center by loading the system starting from zero until 2000 N loads continuously in steps. Linear equation model of each load cell was obtained by applying linear

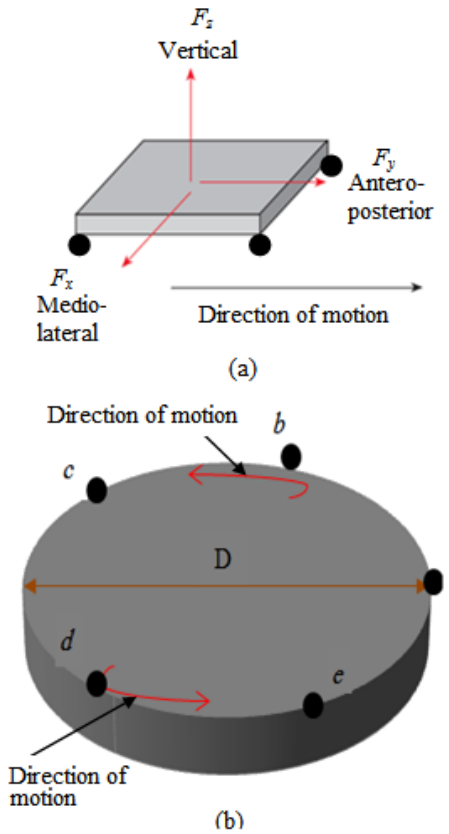


Fig.1. Schematic representation of the force platforms and the involved load cells locations (a) The standard force platform (b) The prposed force platform

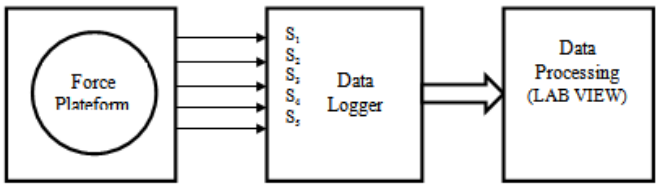


Fig. 2. Block diagram of the proposed technique

Fig. 3. Measurement set up of the fabricated platform

regression of the obtained data to convert each load cell output into Newton unit of the measured force. The load cell output is made to be rated in millivolts per volt (mV/V) and the results are shown in Fig.4 for the five load cells. As shown in the figure, an acceptable linearity of the test for each load cell was presented with error less than 2.5% of the applied load for a full scale of 2000N. Dynamic Calibration is an important issue to test the response of the proposed platform during movement or jumping. In order to achieve this purpose, the platform was subjected to different loads under movement (jumping) conditions and the measured results of forces were compared to the results obtained when the standard force platform (Vernier Platform, USA) available in the Faculty laboratory was subjected to the same loads. This platform is designed to collect data on forces developed during stepping, jumping, and other human-scales actions. Two participants were selected to perform vertical jump and horizontal jump on both platforms with four trials given for each participant to be performed on each one of the platforms. After collecting the data of the measured forces from each platform, a statistical analysis was made to calculate T-statistical value form their obtained average measured values and the standard deviations and the results are listed in Table I. It is clear from the table that the obtained results of indication level of the statistical analysis for both cases is less than the allowable indication value (0.01). The obtained results of this analysis clearly showed that the fabricated platform has performance almost similar to that of the standard platform for the considered cases of vertical jumping and horizontal jumping.

III. Experimental Tests Results

The experimental study and the tests results on the fabricated force platform for the three sport activities (hammer, discus and shot put throw) are presented in this section to access and to evaluate its performance in the real use appropriateness. Six participants, two for each activity, were selected for this study in order to investigate and to access their performances and skills when their respective activities are performed using the fabricated force platform . The details of the information for each one are listed in Table II. Before starting the practical operation validation, the platform is connected to measurement system and wait about 5 minute for load cells signals stabilization.

TABLE I

COMPARISON BETWEEN THE STATISTICAL RESULTS OF THE CALIBRATION TEST PERFORMED ON THE FABRICATED PLATFORM AND THE STANDARD PLATFORM

Statistical value / Variables	Unit	Standard Platform		Proposed Platform		(T) Value	Ind. Level
		Avr	SD	Avr	SD		
Horizontal jump	N	2601	22.61	2598	23.98	1.096	0.000
Vertical Jump	N	3782	49.23	3746	51.87	2.161	0.001

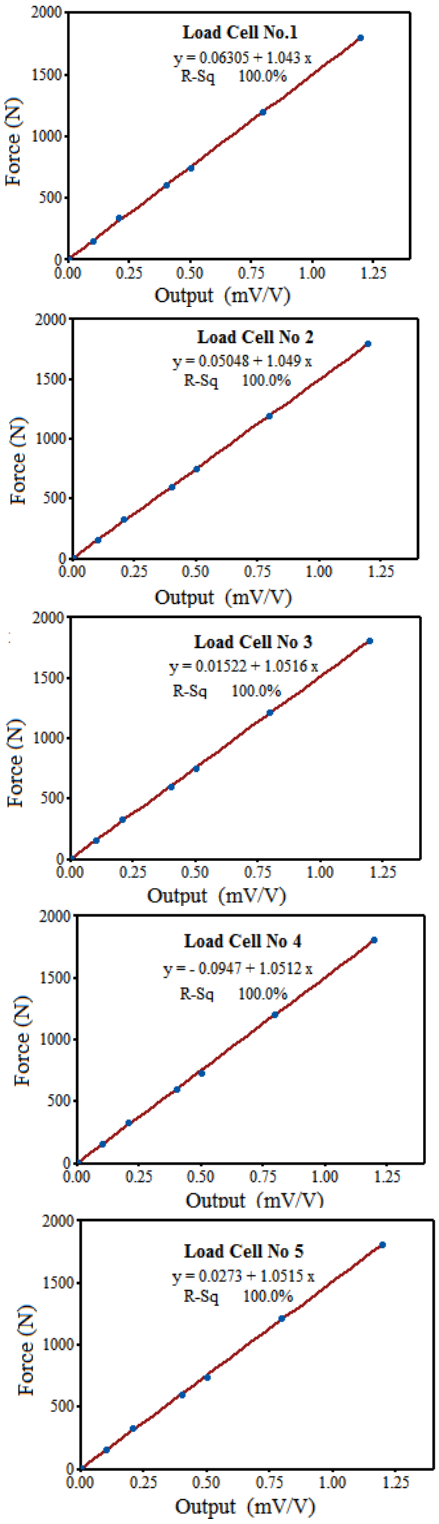


Fig. 4. The force platform calibration results (0 to 2000 N) for each load cell seperatly.

TABLE II.
PARTICIPANTS CHARACTERISATION OF THE SELECTED ATHLETS FOR THE THREE THROWING ACTIVITIES

Activity	Athlete#	Weight (Kg)	Height (cm)
Hammer	#1	115	183
	#2	110	181
Discus	#1	115	189
	#2	114	188
Shot Put	#1	116	185
	#2	115	179

Based on the fact that the force platform is designed to measure continuously with the time the exerted force as well as other parameters of the movement due to rotation, one of the main objectives of this experimental study is to demonstrate the validity of its calibration and operation sensitivity and accuracy for the three activities. The first test was performed for analyzing the measured force with the time for two participants performing hammer throwing activity with their profile listed in Table II and the obtained results is depicted in Fig. 5. From the Figure, it can be observed that a small value of the measured force was obtained for a short period of time. This period represents the static period before the participant starts performing the dynamic period of hammer throw rotation during which the measured force increases rapidly with the time. Also, the sensitivity of the force platform is obvious. It can be seen that a relatively good response of the output was produced by the platform to change from static state to dynamic state for the two participants with short duration of time (measured in the millisecond). The obtained output response gives an indication that the fabricated platform can measure the force with high sensitivity. According to these observed points, Figure. 5 shows a discrepancy in the achieved throwing performances of the two participants (for both static and dynamic periods). The results of the second one showed a static period with a time of 38 s and stable dynamic period with different peaks of the measured force during the time of performing the activity. The first peak of force is with the value of 1230 N while the second and the third peaks are with values of 1355 N and 1421 N respectively before the participant return again to the static period at the end of the activity. For the first participant, a relatively static period with long time (64 S) was obtained compared to the second one results. Also, the obtained results of the first participant during the dynamic period of the activity showed instability in his performance. Figure 6 shows the results of the measured output force of the platform with the time due to rotation of discus throw activity test for two participants with their characteristics listed in Tabe II. It is clear from the figure that the discrepancy in the performances of the two participants is also obvious. Regarding both static and dynamic periods, the second one achieved good performance compared to the first one. For the second one, a static period with 38 s being obtained while the first one a relatively longer time of 45 s being obtained. Also, for the dynamic period, the second one achieved stable rotation performance of the measured force with the time. Force values with peaks of 1289, 1537, 1558, 1466 and 1398 respectively were obtained while the performance of the first one showed instability in the measured forces with the time. Figure 7 illustrates the performances of two participants of the shot put throw activity test for the respective participants with their characteristics listed in Table II where the results of the measured forces with the time for them are plotted. As shown in the Figure, the two participants have different static and dynamic performances of the measured force with the time. Once again, the performance of the second one achieved a stable performance of the measured force with a maximum peak of 1340 N while the first one results showed instability and a discrepancy in the measured force with the time with a maximum peak of 1774 N . Also, it can be seen from the Figure that the time of the static period of the first participant is shorter than the static period of the second one. The obtained results of these activities validate the sensitivity and feasibility of the fabricated force platform in accessing and evaluating as well analyzing performances and skills of athletes in such sport activities. Although the fabricated platform, in the previous sections, is used to measure the force with the time as a result of rotation of movement. One of the most important features of the platform application that it can be used to measure

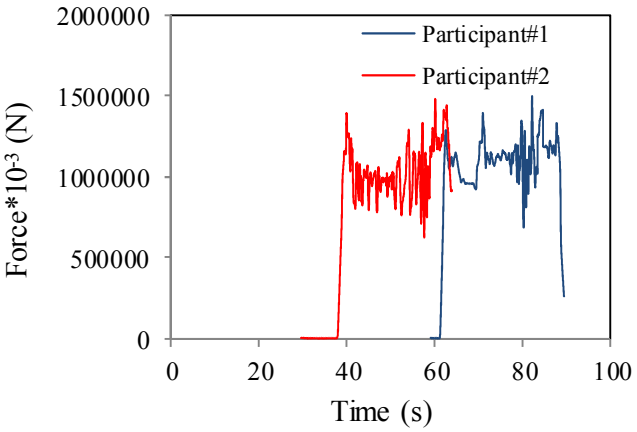


Fig. 5. Variation of the measured force with the time for hammer throwing activity test.

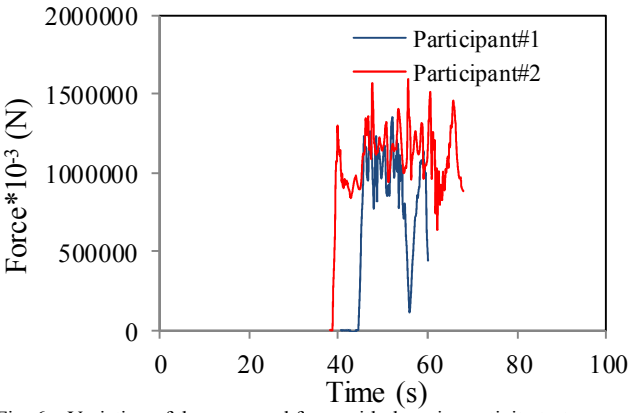


Fig. 6. Variation of the measured force with throwing activity test

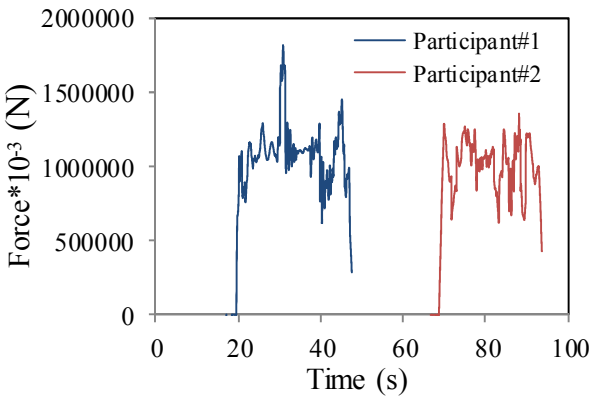


Fig. 7. Variation of the measured platform force with the the time for discus. throwing activity test

different parameters of athletes performances performing these sport activities such as, torque, displacement, and acceleration based on computer vision tracking technique[15] enhanced by camera system research work and this will be appear in the near future.

IV. CONCLUSIONS

In this paper, a new aspect of development and utilization of force platforms have been presented. A ring-shaped force platform has been designed, fabricated, and employed to measure and analyze force due to rotation movements of throwing sport activities, which cannot be realized by using the existing standard force platforms. Measurements of force with the time for various throw sport activities performed using LabView software as virtual instrument have shown the feasibility, reliability and good prospect of performance regarding measurement sensitivity and accuracy. The main feature of the fabricated platform is that its operation can be extended to measure and to measure and analyze different parameters of athletes during performing such activities including torque, moment and displacement using computer vision and tracking system due to the reliability of the developed software. Based on the obtained results of force measurement with the time and analysing athletes performances of the respective throwing sport activities, the designed platform can be served as a technical tool for the scientific and academic community applications, including indoor and in the field laboratories for teaching and monitoring athletes performance progress during training and developing exercises. Further research work such as mathematical sensitivity and accuracy analysis and algorithms improvement should be considered.

REFERENCES

[1] A. Scorza, C. Massaroni, F. Orsini , C. D'Anna , S. Conforto , S. Silvestri and S. A. Sciuto**, "** A Review on Methods and Devices for Force Platforms Calibration in Medical Applications," *Journal of Engineering Science and Technology Review*, Vol. 11 No. 1, pp. 10 – 18, Feb. , 2018.

[2] F. Orsini, A. Rossi ., F. Botta and L. Scalise ," A Case Study on the Characterization of Whole Body Vibration Platforms for Medical Applications," *2018 IEEE International Symposium on Medical Measurements and Applications*, (MeMeA), June 2018.

[3] J. Douglas, S. Pearson, A. Ross, and M. Mcguigen," Kinetic Determinants of Reactive Strength in Highly Trained Sprint Athletes," *J. Strength Cond. Res.*Vol. 32,No.6 , pp. 1562–1570, June, 2018,

[4] I. Loturco, R. Kobal, K. Kitamura, V. Fernandes, N. Moura, F. Siqueira, C. C. Cal Abad, L. A. Pereira, " Predictive Factors of Elite Sprint Performance: Influences of Muscle Mechanical Properties and Functional Parameters," *J. Strength Cond. Res*., Vol. 33, No. 4, , pp. 974-986, April, 2019.

[5] J. B. Morin, P. Edouard, P. Samozino,"Technical ability of force application as a determinant factor of sprint performance,". Med. Sci. Sports Exerc., Vol. 43,No. 9, pp. 1680–1688, Sep., 2011,.

[6] G. Rabita, S. Dorel, J. Slawinski, E. Saez-de-Villarreal, A. Couturier, P Samozino, J. B. Morin, " Sprint mechanics in world- class athletes: A new insight into the limits of human locomotion," Scand. J. Med.Sci. Sports, Vol. 25, No. 5, pp. 583–594, , Feb., 2015.

[7] J. B. Morin, M. Bourdin, P. Edouard, N. Peyrot, P. Samozino and J. R.Lacour," Mechanical determinants of 100-m sprint running performance," *Eur. J. Appl. Physiol*., Vol. 112**,** No.11, pp. 3921–3930, March, 2012,.

[8] M. Coh, and K. Mackala, " Differences Between the Elite and Subelite Sprinters in Kinematic and Dynamic Determination of Countermovement Jump and Drop Jump," *Journal of strength and conditioning research*, ,Vol. 27, No 11, pp. 3021-3027,Feb., 2013.

[9] P. Kerry," Use of force platforms in physics and sport", *Physics Education*, IOP Publishing Ltd, pp. 18-22, 2001.

[10] I. Loturco , L. A. Pereira and R. Kobal R., *et all* " Portable Force Plates: A Viable and Practical Alternative to Rapidly and Accurately Monitor Elite Sprint Performance", *Sports*, Vol. 6, No. 3, pp.61-69, July, 2018.

[11] P. Jiménez-Reyes, P. Samozino , F. Pareja-Blanco et all, " Validity of a Simple Method for Measuring Force-Velocity-Power Profile in Countermovement Jump," *International Journal of Sports Physiology and Performance,*Vol. 12**,** No.1, pp. 36-43, March, 2016 .

[12] L. Mourão , K. de Jesus , N. Viriato, R. J. Fernandes, J. Paulo J., and M. A. Vaz," Design and construction of a 3D force plate prototype for developing an instrumented swimming start block," *Journal of Biomedical Engineering and Informatics*, **2**, 2, pp. 99-108, Sep,. 2016.

[13] A. Scorza, C. Massaroni , F. Orsini, C. D. Anna, S. Conforto, S. Silvestri and S. A. Sciuto," A Review on Methods and Devices for Force Platforms Calibration in Medical Applications," *Journal of Engineering Science and Technology Review*, Vol. 11, No. 1, pp.10 – 18, , Dec., 2017,

[14] G. Paolini , U. Della Croce, P. O. Riley, F. K. Newton and D. Casey Kerrigan," Testing of a Tri- Instrumented-treadmill Unit for Kinetic Analysis of Locomotion Tasks in Static and Dynamic Loading Conditions," *Med Eng Phys*. Vol. 29,No. 3, pp. 404-411, May, 2007.

[15] M. G. Silva, P. V. Sarmet and H. M. Rocha, " Development of a low cost force platform for biomechanical parameters analysis," *Research on Biomedical Engineering*, **33**, 3, p. 259-268, Sep. 2017.

Increasing Autonomy of I/O Cores in Heterogeneous FPGA-SoCs for Data Acquisition Applications

Robert Drehmel
Technische Universität Berlin
Straße des 17. Juni 135
10623 Berlin, Germany
E-mail: drehmel@campus.tu-berlin.de

Markus Hefele
Technische Universität München
Arcisstraße 21
80333 Munich, Germany
E-mail: markus.hefele@tum.de

***Abstract*—As the architectures of FPGA-SoCs become more complex, it is increasingly important to find new ways to optimize the utilization of the interconnects that connect the various components like I/O cores, processors, and memories. Here we present a semi-autonomous soft I²C core and a systematic analysis of the performance characteristics of this core in a Linux system. Compared to a conventional I²C core, the devised core used with a new API reduces the average `irq` processor utilization in monitoring applications by 87.28% for SMBus Read Word operations and by 81.59% for SMBus Write Word operations, while maximizing I²C bus throughput and improving temporal predictability of I²C message transfers. The results show that overhead of interrupt processing to periodically fill a TX FIFO can be substantial and that designing an I/O core to be less dependent on a general-purpose processor can improve system performance considerably.**

***Index Terms*—Field programmable gate arrays, System-on-chip, Operating systems, Data acquisition**

I. Introduction

Modern heterogeneous FPGA-SoCs provide a variety of I/O peripherals connected to processors and memory via a maze of interconnects. The operation of conventional I/O cores for busses like I²C [1] involves frequent communication between I/O core and processor. This induces unnecessary overhead for applications where outgoing traffic is independent of incoming traffic, like data acquisition applications that unconditionally collect measurement data. Novel approaches to increasing the autonomy of I/O cores in heterogeneous systems are needed to reduce this overhead. Data acquisition is often performed on embedded systems where minimizing processor usage and energy consumption are of paramount importance.

In this paper, we introduce the semi-autonomous soft I²C core RTI2C (Real-Time I²C) that maximizes I²C bus throughput, improves temporal predictability of I²C message transfers, and decreases processor utilization. The main contributions of this work are an overview of the design of the RTI2C core and an evaluation of its performance characteristics in comparison with two common I²C cores.

The remainder of this paper is organized as follows. Section II reviews related work, Section III presents the design overview, Section IV details the experimental methodology and setup and discusses the evaluation results, and Section V summarizes the paper, offers a conclusion, and suggests future work.

978-1-7281-9615-2/20/$31.00 © 2020 IEEE

II. Related Work

Whether used in robot hands [2], for cell viability measurements [3], or for a variety of power measurement applications [4], [5], [6], there is a vast variety of I²C sensors accessible as slave devices on an I²C bus. Sensors acting as I²C masters have also been proposed [7]. In contrast, research on I²C master devices to control slave I²C device sensors appears to be stagnant. Most modern SoCs have integrated I²C peripherals (i.e., hard I²C cores). The Intel Cyclone V SoC [8], Arria 10 SoC [9], and Stratix 10 SX [10] include the Synopsys DesignWare APB I²C controller that supports data transfer without CPU involvement using a DMA interface. This is in contrast with the RTI2C core, which uses a local memory (much like the Xilinx Traffic Generator [11] does) and therefore does not tax the interconnects. The Xilinx Zynq-7000 [12] and Zynq UltraScale+ [13] include the Cadence I²C core in revisions `r1p10` and `r114_f0100_final`, respectively. All of these I²C peripherals do not support I²C clock speeds faster than I²C Fast-mode. Soft I²C cores are available in source code or binary form and can be instantiated in an FPGA design. For this paper, we chose one representative each for hard and soft I²C cores: the Cadence I²C peripheral in the Zynq-7000, and the Xilinx AXI IIC Bus Interface v2.0 [14], a soft core that is synthesizable on the Zynq-7000.

III. Design

A. RTI2C Soft IP Core

This subsection gives an explanation of how a common I²C core works, followed by a contrasting description of the RTI2C core. A common I²C core has memory-mapped RX and TX FIFOs. It stores received data in the RX FIFO and a driver stores data to be sent in the TX FIFO. To continuously send I²C messages, the driver has to periodically refill the TX FIFO—to continuously receive I²C messages, the driver has to periodically read (and thereby empty) the RX FIFO. If the fill level of the RX FIFO rises above a threshold or the fill level of the TX FIFO falls below a threshold, the core generates an

interrupt. Upon reception of that interrupt, the driver usually refills the TX FIFO and reads received data from the RX FIFO. The Cadence I²C core operates as outlined above. Its operation is detailed further in [12].

The main idea of the RTI2C core is that repeatedly doing the same task—like continuously reading measurements from the same set of sensors—should not require constant attention from a general-purpose processor. To facilitate autonomous operation, the RTI2C core reads commands from a 1024-byte Block RAM (BRAM) instead of a TX FIFO. A command initiates the transmission of a I²C read or write message on the bus. When the core encounters a command with the `WRAP` bit set, it sets the register containing the BRAM address of the next command to load to a configurable address, the *wrap address*. The BRAM is not memory-mapped, but can be written to through an interface in the memory-mapped register space. The core stores received data in a 1024-byte RX FIFO. It generates an interrupt each time it has executed a configurable number of commands. It stops loading and executing commands when the `STOP_NEXT` bit or the `STOP_WRAP` bit in the control register was set, or when it executed a configurable number of commands. As the names imply, the `STOP_NEXT` bit stops execution after the current command has been executed, whereas the `STOP_WRAP` bit stops execution after the next command with a set `WRAP` bit has been executed.

The RTI2C core supports I²C *Fast-mode* (up to 400 kHz), *Fast-mode Plus* (up to 1 MHz), and *High-speed mode* (up to 3.40 MHz). To fully comply with the timing characteristics of High-speed mode, utilizing active circuitry (e.g., [15]) to pull up the voltage on the bidirectional clock signal is mandatory. However, for sensor chips that do not use *clock-stretching*, the clock signal is unidirectional. Consequently, a push-pull output can be used instead of an open-drain output for the clock signal, avoiding the need for external circuitry.

B. Extension to The Linux I²C Subsystem

The Linux I²C subsystem provides APIs to use I²C cores from user space via device files in `/dev/`. A user space program can open such a device and use the system calls `read()` and `write()` for simple read and write transfers on the I²C bus connected to the I²C core. Alternatively, it can use the system call `ioctl()` to set various options (e.g., the device address for subsequent transfers). The I²C subsystem provides more APIs to start transfers on the bus through the `ioctl()` system call. Besides the command `I2C_SMBUS` that provides an API for transfers on the SMBus protocol [16] level, the command `I2C_RDWR` allows more elaborate I²C bus usage. When an array $M = (m_0, m_1, ..., m_{N-1})$ of I²C message data structures is passed to the `I2C_RDWR` command, the adapter driver is expected to *combine* the messages in such a way that they appear on the bus as the sequence

with S, Sr, and P being start, repeated start and stop sequences, respectively. The adapter driver must copy the data read from an I²C read message into the buffer of the corresponding I²C message data structure. The user program can then process the updated messages containing the data read.

The I²C APIs currently provided by the Linux kernel do not allow taking advantage of hardware that is capable of repeatedly and autonomously processing a set of fixed I²C messages. The following ioctl commands make up the `I2C_CYCLIC` API we propose to leverage that kind of hardware:

- `I2C_CYCLIC_NROUNDS` sets the number of times the I²C message data structure array is processed.
- `I2C_CYCLIC_NROUNDS_IRQ` sets the number of times the I²C message data structure array is processed before the I²C core generates an interrupt request.
- `I2C_CYCLIC_START` configures the I²C core and starts processing.
- `I2C_CYCLIC_FETCH` fetches a given number of I²C message data structures from the kernel.
- `I2C_CYCLIC_STOP` stops the I²C core.

C. RTI2C Kernel Driver

The driver registers itself as an I²C adapter with the I²C subsystem of the Linux kernel. It supports both the `I2C_RDWR` API and the `I2C_CYCLIC` API. When the ioctl commands `I2C_RDWR` or `I2C_CYCLIC_START` are requested, the driver starts the RTI2C core: It converts the array of I²C message data structures into command data structures (similar to the internal format used by the RTI2C core), constructs a program for the core, writes the program into the BRAM through the core's register interface, configures the core by setting its registers to appropriate values, and sets a bit in the control register to start execution of the program.

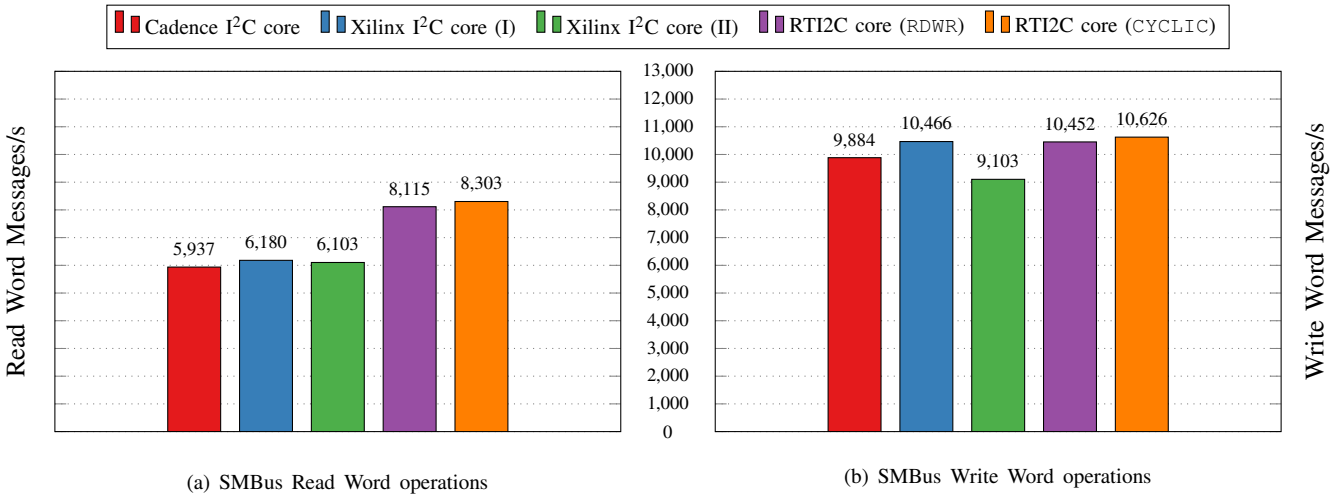
When an interrupt occurs, the driver iterates through the command list from the point where it left of and reads per-command reports from the RX FIFO of the core. For the `I2C_RDWR` API, it copies read data from the reports into the buffers of the corresponding I²C message data structures passed by the user and wakes up the waiting thread. For the `I2C_CYCLIC` API, it stores the read data from the reports in a list of kernel buffers from where it is copied when the user calls the ioctl command `I2C_CYLIC_FETCH`.

To activate High-speed mode, an I²C master device sends a High-speed mode *master code* on the bus. Because this only needs to be done once for each string of combined I²C messages, the kernel driver of the RTI2C core sets the wrap address to the command following the command that triggers sending of the High-speed master code.

IV. Evaluation

A. Experimental Methodology and Setup

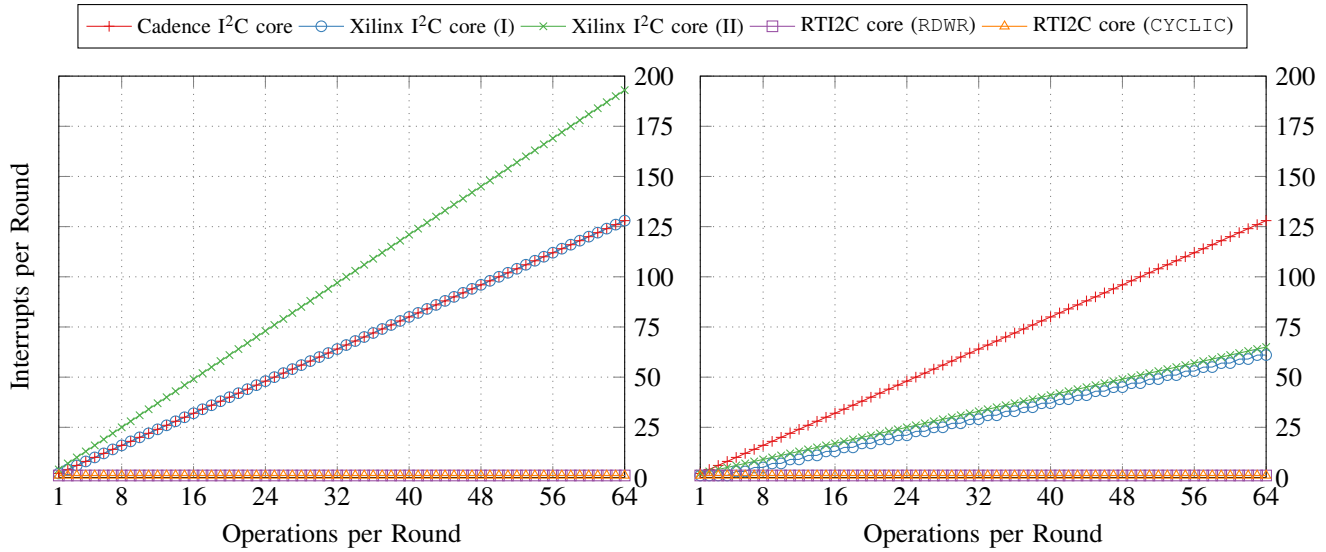
The test methodology aims to simulate a real-life scenario where per "round" N_{op} devices need to be read from or written to using N_{op} SMBus operations. To test Read Word operations, the benchmark tool uses $N_{\text{msg}} = 2N_{\text{op}}$ I²C messages: N_{op} message pairs, each composed of a write message with a data length of one and a read message with a data length of two. To test Write Word operations, it uses $N_{\text{msg}} = N_{\text{op}}$ I²C write messages, each with a data length of 3. The benchmark tool



(a) SMBus Read Word operations

(b) SMBus Write Word operations

Fig. 1. Throughput measurements of the different I²C cores for Read and Write Word operations.



(a) SMBus Read Word operations

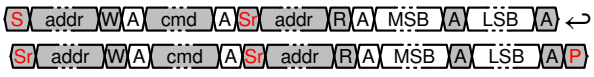
(b) SMBus Write Word operations

Fig. 2. Interrupt measurements of the different I²C cores for Read and Write Word operations.

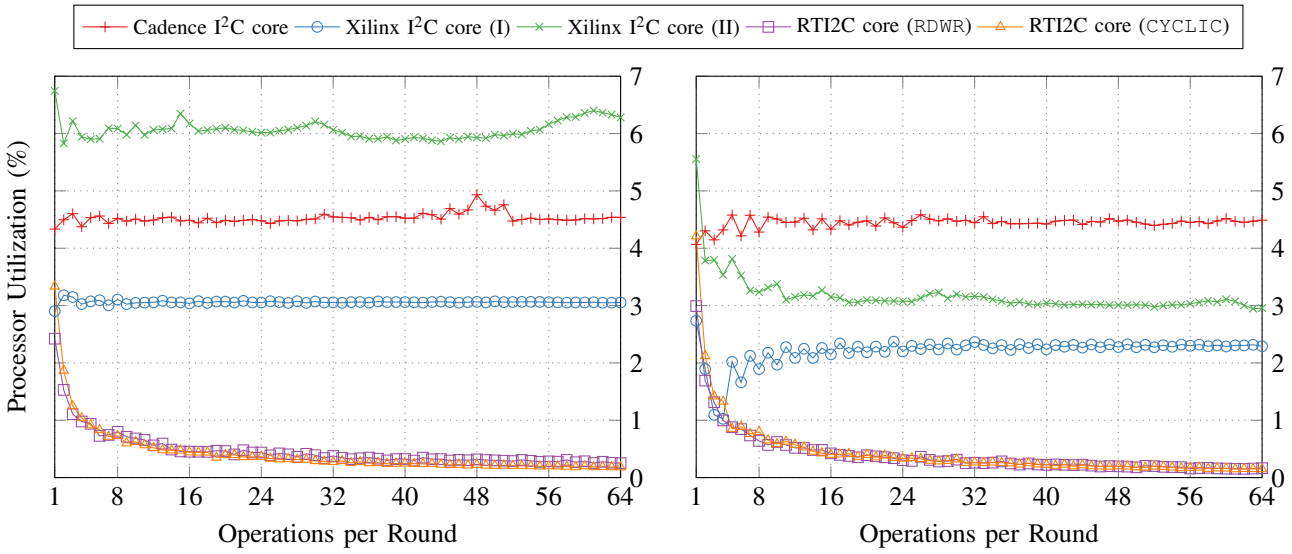
TABLE I
FPGA RESOURCE UTILIZATION

		Xilinx I²C core		RTI2C core	
Resource	Available	Utilization	%	Utilization	%
LUT	17600	380	2.16	818	4.65
LUTRAM	6000	—	—	192	3.20
FF	35200	368	1.05	707	2.01
BRAM	60	—	—	2	3.33
IO	100	2	2.00	3	3.00
BUFG	32	1	3.13	1	3.13

measures throughput, number of interrupts generated from the core, and processor utilization over 10000 rounds and averages the results afterwards. It would ideally only use combined-message API calls. For example, testing Read Word operations in a round with $N_{op} = 2$, the array of I²C message data structures passed to the API would contain 4 elements which are placed on the bus as



Unfortunately, the Linux driver of the Cadence I²C core does not work correctly with combined I²C read messages. Its source code indicates that a hardware quirk is the culprit. The Cadence I²C core Linux driver can currently handle one I²C read message, multiple I²C write messages, or one combined I²C write-read message pair per API call. The Linux driver of



(a) SMBus Read Word operations

(b) SMBus Write Word operations

Fig. 3. Processor utilization (`irq` + `softirq_pc`) of the different I²C cores for Read and Write Word operations.

TABLE II
PROCESSOR UTILIZATION (AVERAGED)

	Read Word					Write Word				
	Cadence	Xilinx (I)	Xilinx (II)	RTI2C (RDWR)	RTI2C (CYCLIC)	Cadence	Xilinx (I)	Xilinx (II)	RTI2C (RDWR)	RTI2C (CYCLIC)
system%	5.56	12.48	7.21	1.25	0.42	3.46	4.56	4.51	1.00	0.27
irq%	4.45	2.98	5.98	0.43	0.38	4.36	2.13	3.09	0.36	0.39
soft_irq%	0.08	0.08	0.08	0.04	0.05	0.08	0.08	0.09	0.05	0.05
Total%	10.09	15.54	13.28	1.72	0.85	7.90	6.77	7.68	1.40	0.71

the Xilinx I²C core exhibited behaviour similar to the Linux driver of the Cadence I²C core when tested with combined I²C messages. We modified the driver of the Xilinx I²C core to be able to process arbitrarily combined messages. Both the Cadence I²C core and the Xilinx I²C core with its original driver were tested with one combined I²C write-read message pair per API call for Read Word operation tests and N_{op} combined I²C write messages per API call for Write Word operation tests. On the other hand, both the Xilinx I²C core with the modified driver and the RTI2C core were tested using N_{op} combined I²C write-read message pairs for Read Word operation tests and N_{op} combined I²C write messages for Write Word operation tests. In the following, "I" and "II" are used as a shorthand notation for the original and modified Linux drivers of the Xilinx I²C core, respectively.

All tests were run on a Digilent Zybo board featuring a Xilinx Zynq-7000 SoC (XC7Z010-1CLG400C, speed grade 1C) [17] and 512 MiB DDR3 RAM. The Analog Devices SSM2603 IC on the Zybo served as the I²C slave. The CPU clock frequency was 650 MHz and the APB bus frequency was 108.30 MHz. Both PL soft cores ran with a clock frequency of 166.67 MHz. All cores were configured to produce a I²C clock frequency of 400 kHz (Fast-mode). The Linux kernel 4.19.0-xilinx was configured with `CONFIG_IRQ_TIME_ACCOUNTING` for more precise IRQ time accounting [18]. Additionally, `I2C_RDWR_IOCTL_MAX_MSGS` was set to 128 to allow a maximum of 128 messages per `I2C_RDWR` API call.

B. Experimental Results

1) Resource Utilization: Table I shows a comparison of FPGA resource utilization of the Xilinx I²C core and the RTI2C core (note that the RTI2C core was configured to use an additional I/O pin to provide the clock both as an open-drain and a push-pull output). The increase in autonomy actualized in the RTI2C core comes at the cost of complexity: The RTI2C core uses more logic resources than the Xilinx I²C core. However, in addition to the fundamental difference in operation, both cores vary in feature set and maximum I²C clock speed.

2) Throughput: Fig. 1 shows the results of the throughput measurements. The Cadence I²C core achieved 5936.65 Read Words/s and 9883.96 Write Words/s. The Xilinx I²C core reached 6180.01 Read Words/s and 10 466.29 Write Words/s

with driver variant I and 6102.56 Read Words/s and 9103.10 Write Words/s with driver variant II. The RTI2C core used with the `I2C_CYCLIC` API was able to execute 8302.68 Read Words/s and 10 625.98 Write Words/s. Using it with the `I2C_RDWR` API resulted in a lower throughput of 8114.80 Read Words/s and 10 452.45 Write Words/s because the hardware was programmed anew for each round. The RTI2C core currently uses a relaxed default protocol timing for S, Sr, and P. Thus, its Write Word throughput when used with the `I2C_RDWR` API was slightly below that of the Xilinx I²C core with driver variant I.

3) Interrupts: Fig. 2 shows the number of interrupts generated. With N_{op} operations in a round, the Cadence I²C core generated $2N_{\mathrm{op}}$ interrupts both for Read Word and Write Word operations. Using driver variant I, the Xilinx I²C core generated $2N_{\mathrm{op}}$ interrupts for Read Word operations and a little less than N_{op} interrupts for Write Word operations. It generated $3N_{\mathrm{op}}+1$ interrupts for Read Word operations and $N_{\mathrm{op}}+1$ interrupts for Write Word operations with driver variant II. The RTI2C core only produced a single interrupt per round, regardless of the number of Read Word operations per round. Note that it is possible to configure the number of rounds per interrupt for the RTI2C core, so even lower interrupt counts may be forced.

4) Processor Utilization: Fig. 3 shows the processor utilization. The percentages shown are the sum of the `irq` and `soft_irq` processor utilization percentages. The averages (e.g., $\frac{1}{64}\Sigma_{N_{\mathrm{op}}=1}^{N_{\mathrm{op}}\leq 64}$ `irq`(N_{op})) of the processor utilization percentages are listed in Table II. The overhead induced by interrupt handling (to fill and empty FIFOs) probably makes up a majority of the processor utilization of the Cadence and Xilinx I²C cores.

V. Conclusions and Future Work

In this paper we have introduced the RTI2C soft I²C core along with its Linux kernel driver and an extension to the Linux I²C subsystem, and we have analyzed specific performance characteristics of the RTI2C core.

In contrast to common TX-FIFO-based I²C cores, the RTI2C core provides predictable execution of I²C messages and maximum throughput when used with the novel `I2C_CYCLIC` API. At the same time, the RTI2C core used with the `I2C_CYCLIC` API reduces the average `irq` processor utilization when processing SMBus Read Word operations by 87.28% and by 81.59% when processing SMBus Write Word operations.

Common I²C cores generate the majority of interrupts for two reasons: to request that the processor provide I²C messages to send and to request that the processor fetch received data. The reduction in processor utilization the RTI2C core achieves stems mainly from eliminating the reason for the first interrupt and reducing the frequency of the second interrupt. First, the processor does not have to repeatedly provide I²C messages to send as the RTI2C core keeps them in its memory. Second, due to the large capacity of the RTI2C core's RX FIFO, interrupts to request its depletion do not have to occur as frequently. A model use case would be to employ the RTI2C core in an embedded system that simultaneously monitors multiple I²C busses—a scenario where minimum processor utilization per I²C core is paramount. Generally, we conclude that giving an I/O core a higher degree of autonomy can prove highly advantageous and suggest that this is true especially for modern FPGA-(MP)SoC architectures that consist of a maze of interdependent components.

Separate analyses of the common I²C cores' interrupts to request RX FIFO emptying and TX FIFO refilling will allow us to shed light on the effect of the larger RX FIFO depth of the RTI2C core. To determine how much impact the design of the RTI2C core has on energy consumption, we plan to compare power and energy measurements of the RTI2C core with those of other I²C cores. Finally, we currently develop an elaborate energy measurement application called Zodiac that utilizes the real-time capabilities of RTI2C to schedule reads from sensors for improved temporal resolution.

Acknowledgements

This research was funded by the *German Academic Scholarship Foundation (Studienstiftung des deutschen Volkes)*.

References

[1] *I²C-bus specification and user manual*, NXP Semiconductors Std., Rev. 6, 2014, UM10204.
[2] T. P. Tomo, W. K. Wong, A. Schmitz, H. Kristanto, A. Sarazin, L. Jamone, S. Somlor, and S. Sugano, "A modular, distributed, soft, 3-axis sensor system for robot hands," in *2016 IEEE-RAS 16th Int. Conf. Humanoid Robots (Humanoids)*, pp. 454–460.
[3] B. Senevirathna, A. Castro, M. Dandin, E. Smela, and P. Abshire, "Lab-on-CMOS capacitance sensor array for real-time cell viability measurements with I²C readout," in *2016 IEEE Int. Symp. Circuits and Systems (ISCAS)*, pp. 2863–2866.
[4] S. Yang, R. A. Shafik, G. V. Merrett, E. Stott, J. M. Levine, J. Davis, and B. M. Al-Hashimi, "Adaptive energy minimization of embedded heterogeneous systems using regression-based learning," in *2015 25th Int. Workshop Power and Timing Modeling, Optimization and Simulation (PATMOS)*, pp. 103–110.
[5] A. Kulkarni, R. Bonamy, and D. Stroobandt, "Power measurements and analysis for dynamic circuit specialization," in *2015 Int. Conf. ReConFigurable Computing and FPGAs (ReConFig)*, pp. 1–6.
[6] K. Railis, V. Tsoutsouras, S. Xydis, and D. Soudris, "Energy profile analysis of Zynq-7000 programmable SoC for embedded medical processing: Study on ECG arrhythmia detection," in *2016 26th Int. Workshop Power and Timing Modeling, Optimization and Simulation (PATMOS)*, pp. 275–282.
[7] W. Andrysiewicz, D. Kościelnik, and M. Miśkowicz, "I2C hardware master serial interface for asynchronous ADCs," in *2015 IEEE Int. Symp. Systems Engineering (ISSE)*, pp. 77–81.
[8] *Cyclone V Hard Processor System Technical Reference Manual*, Intel, 2020, cv_5v4.
[9] *Intel Arria 10 Hard Processor System Technical Reference Manual*, Intel, 2019, a10_5v4.
[10] *Intel Stratix 10 Hard Processor System Technical Reference Manual*, Intel, 2020, s10_5v4.
[11] *AXI Traffic Generator v3.0*, Xilinx, 2019, PG125.
[12] *Zynq-7000 SoC*, Xilinx, 2018, UG585 (v1.12.2).
[13] *Zynq UltraScale+ Device*, Xilinx, 2019, UG1085 (v1.9).
[14] *AXI IIC Bus Interface v2.0*, Xilinx, 2016, PG090.
[15] A. Dragan, A. Enache, A. Negut, A. M. Tache, and G. Brezeanu, "A fast response output buffer for an I²C high speed interface," in *2019 Int. Semiconductor Conf. (CAS)*, pp. 141–144.
[16] *System Management Bus (SMBus) Specification*, SBS Implementers Forum Std., Rev. 2.0, 2000.
[17] *Zynq-7000 SoC Data Sheet: Overview*, Xilinx, 2018, DS190 (v1.11.1).
[18] V. Pallipadi. (2010) Proper kernel irq time accounting. [Online]. Available: https://lwn.net/Articles/405889/

Study on the Biochemical Nanoparticles for Bio-imaging and Molecular Diagnostics of Alzheimer's Disease

Elliot Shi
CRG-NJ
Cresskill, NJ, USA
elliotshi@risegr.com

Amanda Kyung
NVRHS
Demarest, NJ, USA
kyung.amanda@gmail.com

***Abstract* — Alzheimer's disease (AD) is characterized by its step by step deterioration with periods of stability, punctuated with rapid decline in cognitive function. Such deterioration involves the loss of neurons and synapses in the cerebral cortex and certain subcortical regions of the brain.**
This loss causes gross atrophy of the affected regions; which includes degeneration of the temporal and parietal lobe, parts of the frontal cortex and cingulate gyrus, and eventually a reduction in the respective regions of the brain.

In this paper, computational biomedical simulation technology was provided as an alternative method for a solution of AD treatment. A potential solution in applications-related nanotechnology was presented using nano fullerene complexes and EDTA molecules due to their ability to virtually attach large quantities of proton H. Multiple pathways through oxidative stress can produce cell injury in the human brain.
A free-radical chain reaction capable of propagating in space is the major oxidative reaction in bio-membranes in the human brain's nerve system. To find active energy, stability, and efficiency, various clusters of compounds were created and optimization configuration energy was collected. While there was a trend of increase in enthalpy with increasing the chain of the molecules, variation of energy was also observed. The goal of these studies on the role of oxidative stress that plays in the degeneration of nerve cells is how the nano-scaled chemical antioxidants may be useful in the treatment of AD.

Keywords—Alzheimer's Disease, contrast agent, metal oxides, fullerene, computational modeling, Avogadro

I. Introduction

Oxidative stress is the earliest symptom of AD. Excessive accumulation of Reactive Oxygen Species (ROS) in the brain caused by simultaneous dys-functioning of a number of mechanisms results in AD. ROS also results in major damage to proteins and lipids by increasing lipid peroxidation and causing unwanted oxidation of proteins as well as nucleic acids.

Water-soluble derivatives of buckminsterfullerene derivatives are a unique class of compounds with potent antioxidant properties. The onset of Alzheimer's disease (AD) is associated with the presence of neurofibrillary pathology such as amyloid β (Aβ) plaques. Different therapeutic strategies have focused on the inhibition of Aβ aggregate formation; these pathological structures lead to neuronal disorder and cognitive impairment. Fullerene C60 has demonstrated the ability to interact and prevent Aβ fibril development; however, its low solubility and toxicity to cells remain significant problems. In this study, we synthesized, characterized and compared diethyl fullerenemalonates and the corresponding sodium salts, adducts of C60 bearing 1 to 3 diethyl malonyl and disodium malonyl substituents to evaluate the potential inhibitory effect on the aggregation of Aβ and their biocompatibility.

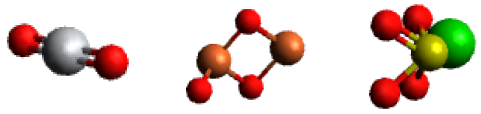
Contrast agents (CA) are drugs that enhance medical imaging of anatomic boundaries for better examination and diagnosis of anomalies in the body, such as tumors. There are many subclasses of contrast agents, often categorized by the imaging technique. For example, iodinated agents are commonly used in computed tomography (CT) or X-ray imaging, and gadolinium agents are commonly used in magnetic resonance imaging (MRI) scans. Iodine atoms in iodinated contrast agents directly interact with the X-ray photon, affecting the attenuation. Gadolinium atoms in contrast agents affect the relaxation of water protons.

Therefore, regions without water, such as bone, show no enhancement in the MRI scan.

II. Metal Oxides And Iodinated Contrast Agents as Contrast Agents

A. Metal Oxides

The ultimate goal of a nano-scaled bioinert material is to efficiently actualize its clinical attributes, one of which is accurately targeting the areas in the body concentrated with the desired product molecule. Following figures show the optimized shapes of the bioinert metal oxides used in the contrast agent.



(a) Titanium oxide (b) Iron oxide (c) Barium Sulfate
Fig. 1 Exceedingly small nanoparticles as positive MRI contrast agent

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Computational analysis can help understand the stability and activity of the Titanium oxide. The graph in Fig.2 illustrates how TiO_2 stabilizes and the steps that it takes. The results show that energy converges at the 7th iteration calculation, which is relatively fast.

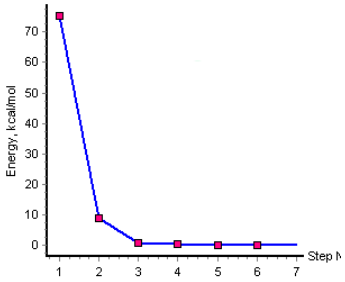


Fig. 2 Optimization convergence of the TiO_2

Gadolinium is a lanthanide element, which has become commonly used in medical treatments. The unique magnetic properties of the gadolinium ion allowed it to be used in a variety of ways, especially in MRI scanning. Following figure shows the optimized shapes of the bioinert gadodiamide used in the contrast agent.

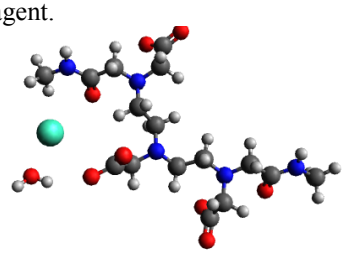


Fig. 3 Model of gadodiamide

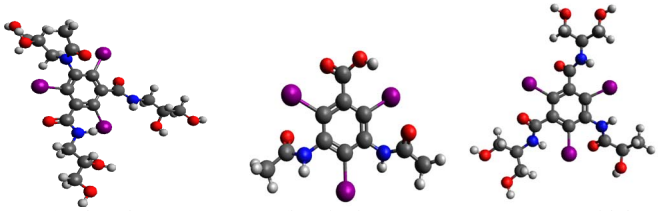
As shown in Table 1, the titanium oxide has the lowest optimization energy and barium sulfate has the highest. The Gadodiamide clusters took the longest to be optimized.

TABLE I. OPTIMIZED ENERGY OF PARAMAGNETIC METALS AND GADODIAMIDE

Compound	Energy (kJ/mol)
Iron oxide	355.73
Titanium oxide	3.30161 * 10e-5
Barium sulfate	424.86
Gadodiamide	321.467

B. Iodinated Contrast Agents

Also, iodinated Contrast Agents have become commonly used in the MRI scanning. Following figure shows the optimized shapes of the bioinert iodinated Contrast Agents used in the contrast agent.



(a) Iohexol (b) Diatrizoate (c) Iopamidol

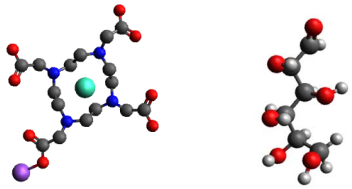
Fig. 4 Iodinated contrast agents

TABLE II. OPTIMIZED ENERGY OF IODINATED CONTRAST AGENTS

Compound	Energy (kJ/mol)
Iohexol	731.689
Diatrizoate	455.739
Iopamidol	656.326

B. Biodegradable Contrast Agents

Biodegradable contrast agents have become popular. Following figure shows the optimized shapes of the biodegradable contrastagents used in the contrast agent.



(a) D-glucose (b) Polysulfide Gd-DOTA

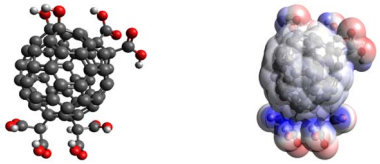
Fig. 5 Biodegradable contrast agents

TABLE III. OPTIMIZED ENERGY OF BIODEGRADABLE CONTRAST AGENTS

Compound	Energy (kJ/mol)
D-glucose	159.299
Polysulfide Gd-DOTA	127.365

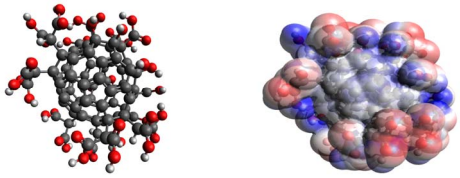
II. SOLUBLE FULLERENES

Literature shows the effectiveness of fulleren-based molecules as MRI contrast agents. Comparing the optimized energies of these functionalized fullerenes leads to a conclusion of molecular safety and stability. In this section, C60 molecules, each doped with OH, COOH, and $C(COOH)_2$ clusters, were modeled in the Avogadro software.



(a) (b)

Fig. 5 (a) 2 OH, 2 COOH, 2 malonic acid on C60
(b) 6 OH, 6COOH, 6 malonic acid on C60 with the electrostatic potential map



(a) (b)

Fig. 6 (a) 6 OH, 6COOH, 6 malonic acid on C60
(b) 6 OH, 6COOH, 6 malonic acid on C60 with the electrostatic potential map using Avogadro software

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

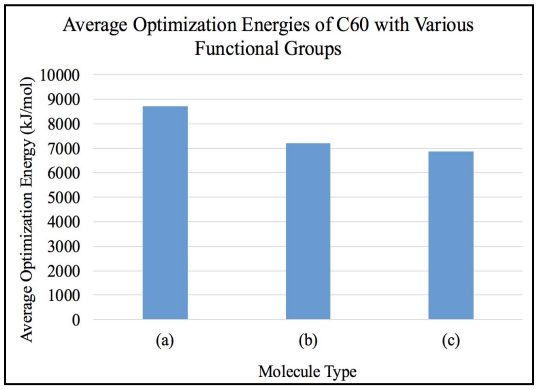


Fig. 7 Average optimization energies of C60 with various functional groups. (a) C60 with 2 OH, 2COOH, 2 malonic acid ($C_3H_4O_4$). (b) C60 with 4 OH, 4COOH, 4 malonic acid ($C_3H_4O_4$). (c) C60 with 6 OH, 6COOH, 6 malonic acid ($C_3H_4O_4$)

III. Functionalized Fullerenes as Contrast Agents

$PC_{61}BA$-(Gd-DO3A) is a functionalized fullerene that is created through synthesis of $PC_{61}BA$ and Gd-DO3A. Literature shows the effectiveness of gadolinium-based molecules as MRI contrast agents.

The Gd-DO3A and other 6 molecules were modeled through Avogadro to check their safety and stability.

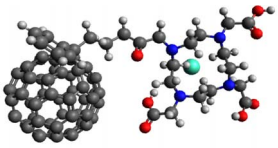


Fig. 8 Gd-DO3A with $PC_{61}BA$.

After modeling the functionalized gadolinium chelates, a table was constructed with the data. Using the energy values, a conclusion on thermodynamic stability of each molecule was made.

TABLE IV. Optimized Energy of Gadolinium Chelates with Ligands

Compound	Energy (kJ/mol)
Gd(DTPA) Magnevist	107.254
Gd(DPTA-BMA) Omniscan	109.196
Gd(BOPTA) MultiHance	191.118
Gd(DO3A-butrol) Gadovist	195.242
Gd(DPTA-BMEA) OptiMARK	131.679
Gd(DOTA) Dotarem	151.421
Gd(HP-DO3A) ProHance	150.370

Table IV shows optimized energy of all the seven Gd functionalized groups. Gd(DTPA) Magnevist has lowest energy, showing that it is the most thermodynamically stable.

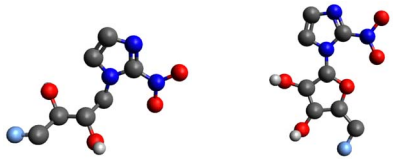
IV. Radiolabels used as Positron Emission Tomography (PET) Hypoxia Tracers

Each of the six PET hypoxia tracers, EF5, FMISO, FETNIM, FAZA, HX4, and Cu-ATSM, are at different validation and evaluation stages.[10] FMISO, first published in 1987, has been thoroughly validated and evaluated. FMISO is used for various tumor types, including head/neck, lung, brain, colon/rectum, prostate, breast, pancreas, and kidneys. FAZA is now used for tumors in head/neck, lung, brain, cervix, colon/rectum, and prostate. A direct comparison of the PET radiotracers is difficult due to some PET tracers being highly specific to certain types of tumors. For example, while EF5, FMISO, FAZA, and Cu-ATSM are used for tumors in the brain, FETNIM and HX4 are not typically used for brain tumors. The specificity of each radiotracer depends on the level of uptake in different tumor types in the body.[5]

With Cu-ATSM, limitations exist due to inter-tumor differences in hypoxia selectivity that prevent the radiotracer from being widely used. [5]

TABLE V. Optimized Energy of Gadolinium Chelates

Compound	Energy (kJ/mol)
EF5	176.459
FMISO	169.195
FETNIM	187.285
FAZA	321.893
HX4	364.444
Cu-ATSM	132.602



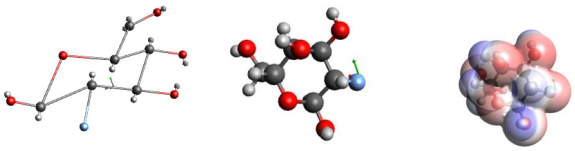
(a) FETNIM (b) FAZA

Fig. 9. Structure of PET hypoxia tracers

A molecule will be stable if there are more attractive forces and less repulsive forces. Attractive forces lower the potential energy of the molecule, and repulsive forces increase the potential energy of the molecule.

V. Electronegativity and Reactivity

18F-FDG is a readily transportable radiotracer commonly used in radiolabeling as it phosphorylates into FDG-6-phosphate after being injected into cells.

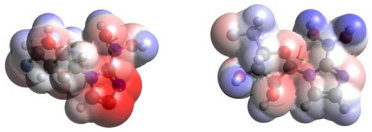


(a) 18F-FDG (b) Optimized 18F-FDG (c) EP map

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Fig. 10 18F-FDG is more reactive than the normal unlabeled Glucose and readily transportable radiotracer

Additionally, the fluorine in the molecule, with an electronegativity of 4.0, is the element with the greatest electronegativity and the most reactive of all nonmetals. Extremes in electronegativity causes increased reactivity.



(a) FETNIM (c) FAZA

Fig. 11 Electrostatic potential maps of radiotracers and corresponding optimization energy and electronegativity

The IA Alkaline elements are the most active metals as they are the family of elements with the lowest electronegativity and most activity.

VI. Discussions and Conclusions

Contrast agents have important implications in medical imaging, giving physicians the ability to visualize diseases and injuries in the human body. Contrast agents can be studied through empirical and computational research, to construct more safe and stable alternatives that have minimal side effects. This research examines the potential of functionalized fullerenes, gadolinium chelates and other nanoparticles as contrast agents by studying their thermodynamic stability.

In this research, Avogadro and Chemcraft were used to model and optimize the nanoscale molecular compounds. The optimized energy was compared and analyzed to deduce the safety and stability of such molecules. In addition, the Density Functional Theory (DFT) was used to study the electron properties of the compounds.

The theoretical and stereo-chemical structure of each feasible nano-scaled metal oxides, fullerene derivatives, and compounds for PET were studied in this project. Based on the predicted stability of each molecule, it can be determined which compound can be used more efficiently to assess thermodynamic stability. Using quantum chemical calculations, it was determined whether molecules can or cannot be formed, as well as can or cannot be used as a biochemical compound for contrast agent.

References

[1] Wharton, T., & Wilson, L. J. (2002). Highly-Iodinated Fullerene as a Contrast Agent For X-ray Imaging. Bioorganic & Medicinal Chemistry, 10(11), 3545-3554. doi:10.1016/s0968-0896(02)00252-3

[2] J Zhang, Y., Zou, T., Guan, M., Zhen, M., Chen, D., Guan, X., Shu, C. (2016). Synergistic Effect of Human Serum Albumin and Fullerene on Gd-DO3A for Tumor-Targeting Imaging. ACS Applied Materials & Interfaces, 8(18), 11246-11254. doi:10.1021/acsami.5b12848

[3] https://www.radiologysolutions.bayer.com/products/mr/contrast/gadavist/

[4] McMurry, and Randall B. Lauffer (1999). Gadolinium (III) Chelates as MRI Contrasting Agents: Structure, Dynamics, and Applications. EPIX Medical, Inc., 2293-2296, 2301-2303.

[5] [1] Hammond, Ester & Asselin, M-C & Forster, Duncan & O'Connor, J.P.B. & Senra, J.M. & Williams, K.J. (2014). The Meaning, Measurement and Modification of Hypoxia in the Laboratory and the Clinic. Clinical oncology (Royal College of Radiologists (Great Britain)). 26. 10.1016/j.clon.2014.02.002.

[6] https://oncohemakey.com/nanoparticles-for-cancer-imaging-and-therapy/

[7] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5695152/

[8] https://en.wikipedia.org/wiki/Fludeoxyglucose_(18F)

[9] https://en.wikipedia.org/wiki/Positron_emission

[10] https://www.ncbi.nlm.nih.gov/pubmed/26203928

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Dielectric characterization of Borofloat 33 and some classic glasses by the coaxial cable method

Abdoulaye SISSOKO
University of Nantes
CNRS, IETR-UMR 6164
Nantes, France
abdoulaye.sissoko@etu.univ-nantes.fr

Anne CHOUSSEAUD
University of Nantes
CNRS, IETR-UMR 6164
Nantes, France
Anne.Chousseaud@univ-nantes.fr

Tchanguiz RAZBAN
University of Nantes
CNRS, IETR-UMR 6164
Nantes, France
Tchanguiz.Razban-Haghighi@univ-nantes.fr

Badié DIOURTÉ
Université des Sciences, des Technique et des Technologies de
Bamako, Mali
badiediourte@yahoo.fr

***Abstract*—In this study, we made the technical qualification of borofloat 33 glass and also other much cheaper conventional glasses, associated with solid metallization by a broadband characterization method based on the S-parameters. This made it possible to extract the electromagnetic parameters (dielectric permittivity, tangent loss) necessary for the design and construction of low visual impact antennas. The data of these measurement results are processed. The values of S_{11}, S_{33}, S_{12}, and S_{34} are then extracted. The values of the loss tangent and the dielectric constant are thus measured. We have obtained for the classical glass ε_r = 6.5 and tanδ = 0.0206; for the microscope glass ε_r = 6.15 and tanδ = 0.021 and at the end for the Borofloat glass ε_r = 3.9 and tanδ = 0.012.**

Keywords—glass, dielectric permittivity, tangent loss, S parameters.

I. Introduction:

Antennas are more and more present on a daily basis with the development of mobile communications and connected objects. Two major challenges are to be taken up, in particular the integration of antennas into communication devices and the reduction of the visual impact of antennas on the public. Hence the growing need to produce antennas with low visual impact and at a bearable cost, which will allow implementation on glazed surfaces (glazing of buildings, vehicles, smartphones, etc.) and installation on historic sites and shopping centers to improve network coverage but also for greater discretion to the public thanks to their appearances. Planar antennas are a category of antennas which, in essence, are discrete (due to their form factor, their weight, etc.) [1] and which also allow the implementation of miniaturization techniques. To design planar antennas, you have to rely on both dielectric and conductive materials. If we are primarily interested in dielectric materials, we find two main categories, glasses, and so-called "plastic" materials. We start with materials whose transparency properties are guaranteed but for which no estimate has been made of the value of fundamental electromagnetic parameters for the design of antennas, namely the relative permittivity and the loss tangent of the dielectric substrate.

Electromagnetic characterization of materials has become increasingly important because it helps to properly size the antennas and, therefore, to predict their accurate frequency behaviors before manufacture [2]. Methods of measuring materials, although numerous, are generally divided into two main groups: broadband (transmission lines,) and narrowband (resonant cavity, waveguide, free space).

Methods that use a cavity resonator can characterize the material at a single frequency or certain discrete values of frequencies. They are divided into two groups:

- The first is the dielectric resonance technique in which the dielectric material itself acts as a resonant element, but it is limited only to low loss samples.
- The second is the perturbation technique, the presence of a small sample in a resonant cavity causes a field disturbance in the latter, and thus a resonant frequency shift [3]. This

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

technique is suitable for samples with low loss and medium loss.

Reflection transmission methods allow measurement over a wide range of frequencies. The classification of materials is done by determining the reflection and transmission coefficients resulting from characteristic changes in impedance and wave speed.

II. Presentation of the technology used

The electromagnetic characterization method used is a broadband measurement method. In this method, only the fundamental waveguide mode (TEM mode in the coaxial lines, quasi-TEM mode in transmission lines and TE_{10} mode in the waveguides) is assumed to propagate. This technique takes into account the fact that the dielectric substrate will be used in microstrip technology for the design of the antenna, ie with metallization above / below.

A measurement using this method involves placing a sample in a section of waveguide line or coaxial line and measuring the diffusion parameters of the two-port complex with a vector network analyzer (VNA) (figure1). The method comprises measuring the reflected (S_{11} or S_{22}) and transmitted (S_{21} or S_{12}) parameters. The relevant diffusion parameters are closely related to the complex permittivity and the permeability of the material by equations. For an accurate dielectric measurement, the maximum electric field is required in the sample. Figure 1 illustrates the typical measurement setup with a coaxial cable.

The advantages of the transmission/reflection line method include:

- coaxial lines and waveguides are commonly used to measure samples with medium or high loss;
- it can be used to determine the permittivity, the permeability and the loss tangent of the material tested.

Some disadvantages of the transmission/reflection line method also exist:

- the accuracy of the measurements is limited by the effects of the air gap.
- it is limited to low precision when the length of the sample is the multiple of half a wavelength of the material.
- the preparation of samples is relatively difficult because of its different stages of realization

The propagation constant γ is related to the attenuation coefficient α and the phase constant β by the relation $\gamma = \alpha + j\beta$. For a dielectric material, the propagation coefficient of a wave in a transmission line is related to the complex permittivity of the filler material through the relation [4]:

$$\gamma = j\sqrt{\frac{\omega^2 \mu_r \varepsilon_r}{c^2} - \left(\frac{2\pi}{\lambda_c}\right)^2} \qquad (1)$$

Where ω is the angular frequency, μ_r is the permeability of the material which is equal to 1, c is the speed of light and, λc is the cutoff wavelength of the transmission line. In case of a coaxial transmission line supporting TEM propagation the cutoff wavelength is taken to be equal to infinity. $\varepsilon_r = \varepsilon_r' + j\varepsilon_r''$; is the complex permittivity of the material filling the line where, ε_r' is the dielectric constant and ε_r'' is the dielectric loss of the medium. From equation 1 it becomes possible to extract the complex permittivity of the material filling the transmission line.

$$\varepsilon_r' = \left(\frac{c}{\omega}\right)^2 \left[\left(\frac{2\pi}{\lambda_c}\right)^2 - \alpha^2 + \beta^2\right] \qquad (2) \text{ and}$$

$$\varepsilon_r'' = 2\alpha\beta\left(\frac{c}{\omega}\right)^2 \qquad (3)$$

This technique is relatively easy to be implemented both in measurement and in the process of extracting parameters.

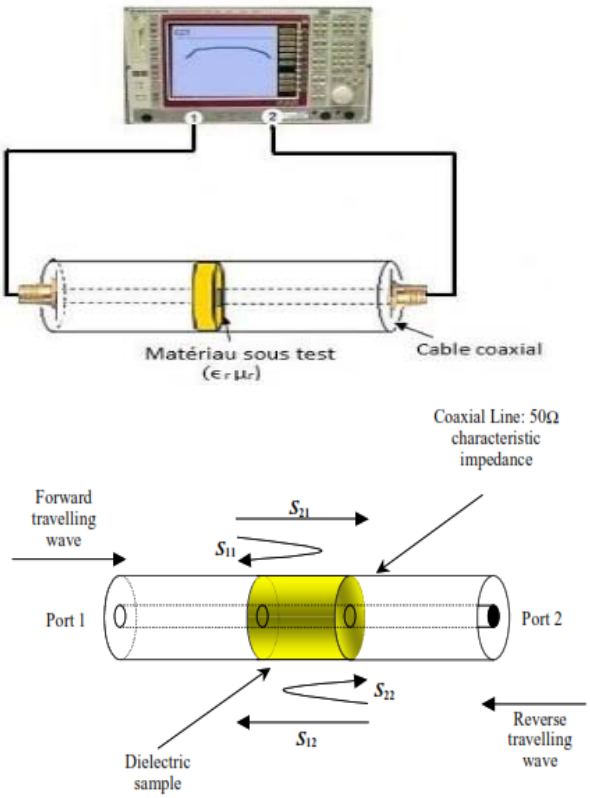


Figure 1. Measurement technique of a dielectric substrate taking into account the application to which it is intended i.e. antenna design [5]

For TEM mode, the complex relative permeability and permittivity can be found as [6]:

$$\varepsilon_r = \frac{Z_0 \gamma \lambda}{j Z_s 2\pi} \quad (4) \qquad \mu_r = \frac{Z_s \gamma \lambda}{j Z_0 2\pi} \quad (5)$$

Where Zs is the characteristic impedance of the sample, Z0 is the characteristic impedance of air for the same dimensions, λ is the wavelength of free space, and γ is the propagation constant which is determined by function of the S parameters as follows:

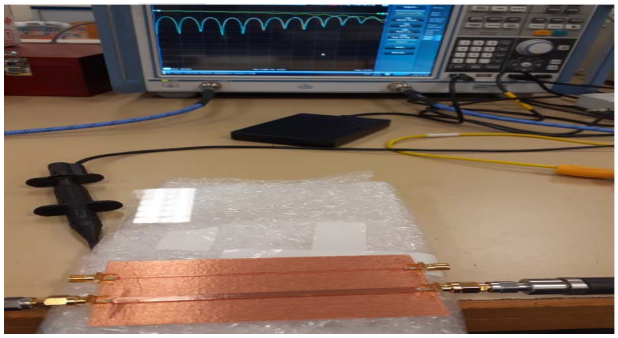
$$\gamma l = cosh^{-1}\left(\frac{1-S_{11}^2+S_{21}}{2S_{21}}\right) \quad (6)$$

Where l is the thickness of the sample

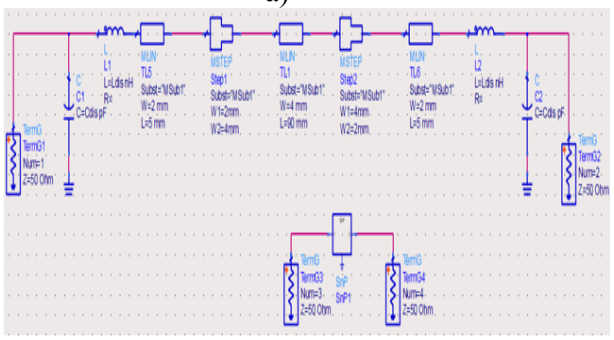
III. Materials and methods

This technique was first applied to two conventional glasses (typically microscope glass and bulk glass) and which transparency is about 90% in the visible range and then to borofloat 33. The idea is to work with transmission lines of impedance clearly different from 50Ohms so as to generate sufficiently clear mismatches to be able then to associate a model with the measurement and to identify the values of ε_r and tan_δ parameters which make it possible to fitter the custom-made model. On each glass, we put two metallization with 35um thick copper and four 50 Ohm ports. The upper metallization comprises two lines and the lower one constitutes the ground plane. One of them has an impedance greater than 50 Ohm with a width of 0.4mm and the second line has an impedance less than 50 Ohm with a width of 8mm.

The measurements were made on a network analyzer (VNA) after calibration. A first calibration step is also required various but well-known terminations (such as open load, short circuit, or 50 ohms for example). The practical values of the reflection and transmission parameters S are thus measured for each of the lines on the VNA. The material under test is modeled on a 1D simulation software ADS (Advanced Design System). The equivalent model is designed according to the shape of the lines produced. At both ports, we have the 50 Ohm ports which are modeled by two 50 Ohm generators each linked to a capacitor and an inductor. Then come two 50 Ohm lines which are subsequently connected to the lines of impedance lower or higher than 50 Ohm by a step. The electromagnetic parameters and values of components used in the port models are adjusted to make the theoretical and practical response curves coincide. The values of the loss tangent and the dielectric constant are thus extracted.



a)



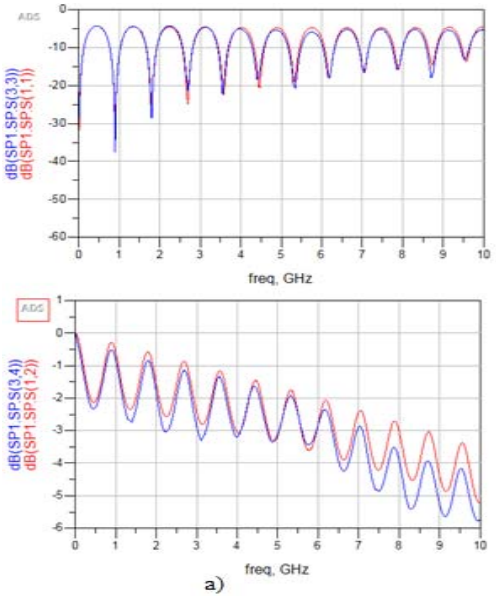
b)

Figure 2: a) measurement configuration on VNA
b) Modeling on ADS

IV. Simulation Results

A. Classic glass:

A classic very low-cost glass has been characterized, the measurement results are below (figure 3). The curves in blue represent the results of practical measurements made on the VNA, those in red are those of modeling with ADS.



a)

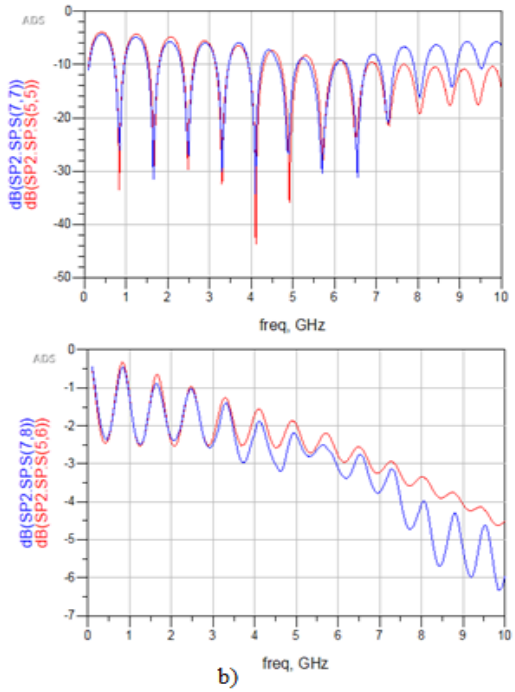


Figure 3: Frequency responses of conventional glass. a) the impedance line Z <50 Ohm (8mm) b) the line of impedance Z> 50 Ohm (0.4mm)

The model is good up to 8GHz since it allows a good correlation with the measurement. The extraction of electromagnetic parameters such as relative permittivity and loss tangent are sufficiently precise to then allow a quality antenna design in this frequency band. We have obtained a dielectric permittivity of 6.5 and a loss tangent of 2.06 10^{-2}. These values remain high for antenna design, specifically ε_r and will have an unfavorable effect on the radiating performance. Conversely, this helps miniaturization effect.

B. Microscope glass

As with classical glass, we made the same simulations, the following results were obtained with the microscope glass: relative dielectric constant ε_r = 6.15 and loss tangent tanδ = 0.021. The resonance curves are shown in figure 4.

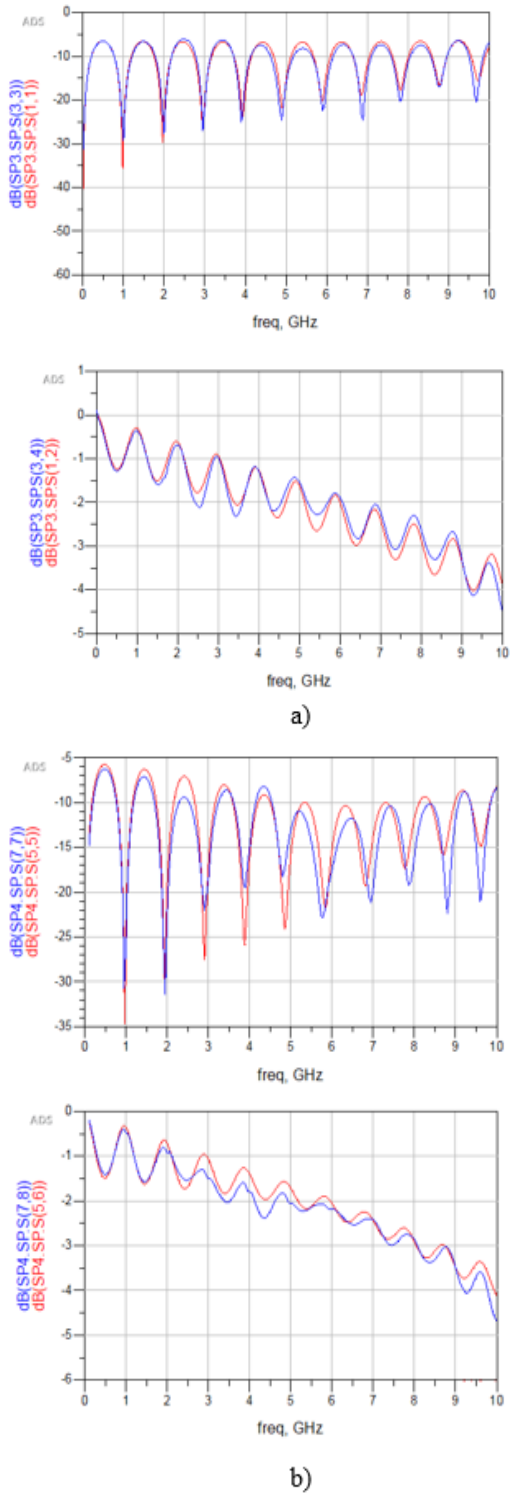


Figure 4: Frequency responses of the microscope glass a) the impedance line Z <50 Ohm (8mm) b)) the impedance line Z> 50 Ohm (0.4mm

The dielectric characteristics of this glass are close to that of conventional glass, as are the resonance curves of the S parameters. It can, therefore, be used in the same devices as conventional glass.

C. *Borofloat glass 33*

Figure 5 shows the resonance curves of borofloat 33. By the same method, we obtained $\varepsilon_r = 3.9$ as the value of the dielectric permittivity and $\tan\delta = 0.012$ as the value of the loss tangent.

Boroflaot 33 has a lower dielectric permittivity and loss tangent than the first two glasses. The resonance curves of the S parameters offer good performance in the 0.1GHz - 7GHz band.

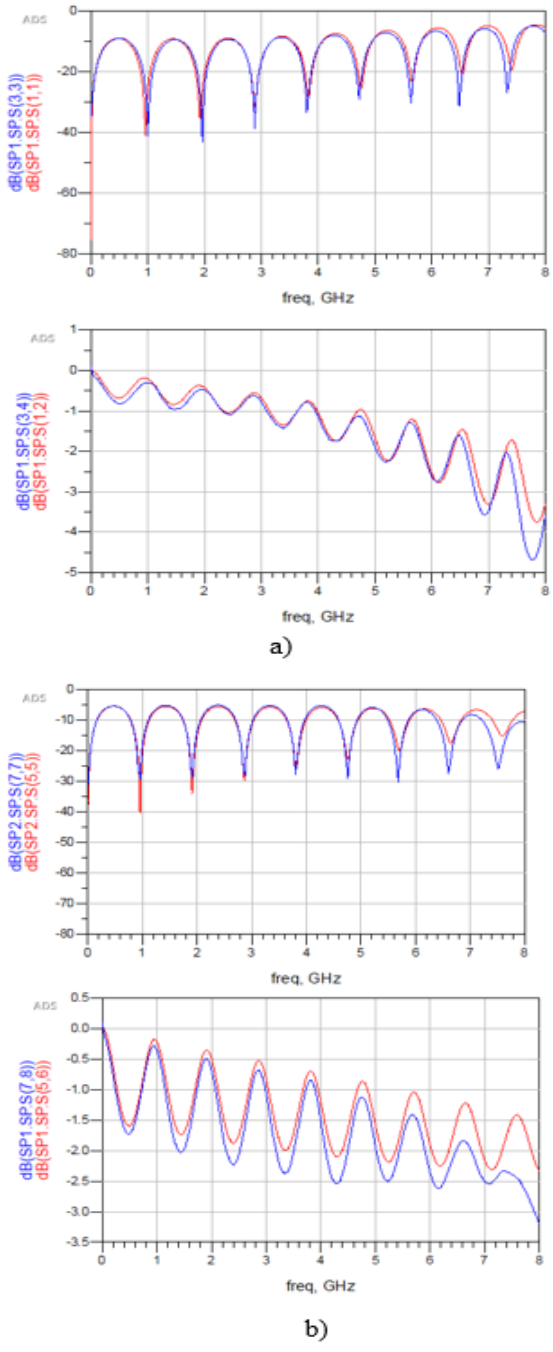


Figure 5: Frequency responses of borofloat 33 glass. a) the impedance line Z<50 Ohm (8mm) b) the impedance line Z> 50 Ohm (0.4mm

V. Conclusion

This article allowed us to find the dielectric parameters of a few glasses for the realization of antennas with low visual impact and low cost of production. It presents on the one hand the electromagnetic characterization method that was used to characterize glass substrates used in the design of transparent antennas. This method has several advantages. It is broadband and thus makes it possible to extract the relative permittivity and tangent parameters of the loss angle of the dielectric substrate over a frequency band of several GHz. This extraction can be done on a more or less important frequency band depending on the precision of the model used. In addition, the extraction of the parameters of the substrate is done in the configuration/technology which will be retained subsequently for the design of the antennas. In this sense, it is almost more of a characterization of the technology used to design the antennas than a characterization of the substrate alone as is the case with some characterization techniques. On the other hand, this article presents the relative permittivity and loss tangent values that can be expected from this type of material, whether on very standard glasses or on glasses intended for RF applications. Among these glasses, borofloat 33 is the best alternative with its low dielectric permittivity and its reasonable loss tangent for the realization of an antenna with the low visual impact a low cost.

References

[1] Ung Hee Park, Haeng Sook Noh, Seong Ho Son, Kyong Hee Lee, and Soon Ik Jeon. A novel mobile antenna for ku-band satellite communications. ETRI Journal, vol. 27(no. 3) :pp. 243–249, June 2005

[2] J.Wang, J. Tao F. Benmohamed “An Efficient Numerical Method to Analyze Discontinuity of Coaxial Cable” International Applied Computational Electromagnetics Society Symposium - Italy (ACES) 2 pages, March 2017

[3] J. Sheen, "Amendment of cavity perturbation technique for loss tangent measurement at microwave frequencies," Journal of applied physics, vol. 102, p. 014102, July2007.

[4] M. B. B. Kashif Saeed, Ian C. Hunter Ed, " Planar Microwave Sensors for Complex Permittivity Characterization of Materials a Their Applications " Applied Measurement Systems. In-Tech, pp 319-350, February 2012.

[5] Andrew P. Gregory and Robert N. Clarke « A Review of RF and Microwave Techniques for Dielectric Measurements on Polar Liquids » 5 pages, IEEE Transactions on Dielectrics and Electrical Insulation Vol. 13, No. 4; August 2006

6] J. Shenhui, D. Ding, J. Quanxing "Measurement of electromagnetic properties of materials using transmission/reflection method in coaxial line", presented at the The 3rd Asia pacific Conference on Environmental Electromagnetics CEEM China, pp 590-595, November 2003

ACKNOWLEDGMENT

The research is supported by the government of Mali as part of the Training of Trainers Program.

Object detection and classification by cascade object training

Ahmed Masud Chowdhury
Department of Electrical & Computer Engineering
Indiana University – Purdue University Indianapolis
Indiana 46202, United States
amasudch@purdue.edu

Jafrina Jabin
Department of Electrical & Computer Engineering
North South University
Dhaka 1229, Bangladesh
jafrina.jabin@northsouth.edu

Erteza Tawsif Efaz*
Department of Electrical & Electronic Engineering
Ahsanullah University of Science & Technology
Dhaka 1208, Bangladesh
ertezatawsif@gmail.com

Md Ehtesham Adnan
Department of Electrical & Computer Engineering
North South University
Dhaka 1229, Bangladesh
ehtesham.adnan@northsouth.edu

Ashfia Binte Habib
Department of Electrical & Computer Engineering
North South University
Dhaka 1229, Bangladesh
ashfia.habib@northsouth.edu

***Abstract*— Computer Vision (CV) has become ubiquitous in smart systems for detecting and labeling objects, starting from social media platforms to autonomous vehicles. It requires extensive computation and image processing. In this paper, a model is processed and used to detect various colored cups with saucers from a set of different objects. The system is trained using Cascade Trainer Graphical User Interface (GUI), and the testing is done utilizing MATLAB, discussed in detail. Finally, the model is tested for its efficacy on the S32V234 Evaluation Board (EVB). Our proposed system accomplished its goal by identifying and tagging the objects of interest with maximum possible accuracy.**

Keywords*— *MATLAB, S32V234, object detection, computer vision, HAAR Cascade Classifier

I. INTRODUCTION

Object detection and classification is a core part of Computer Vision (CV). In autonomous systems, image detection is one of the features that can make autonomous vehicles smart. Generally, a system that can identify any object through CV can be considered as a smart system [1-3].

Detection methodologies have progressed on a good note by the exceptional work of Viola & Jones by introducing a speedy object identification theme supporting a boosted cascade of straightforward feature classifiers [4, 5]. Due to the simplicity and effectiveness of HARR-type selections, this system can process pictures extraordinarily quickly with the achievement of high detection rates, which may fulfill the minimum requirements of many feasible applications such as, human face recognition in regards to Human-machine Interaction (HMI). However, abundant work is required before automatic object identification technologies can perform like humans [6].

Viola-Jones' technique of object detection is one amongst the foremost utilized in the resolution of identifying tasks. But this detection pace is achieved including a few incorrect allowances of unspecified objects, although the refusal error for specified objects is considerably low [7, 8]. The rationale is that in essence, Viola-Jones' refusal cascade of classifiers is integrated by AdaBoost to make a robust classifier with every node being a collection of weak classifiers employing HAAR-type selections. Here, the cascaded classifier is a supervised classifier, in which sub-windows of multi-resolutions of an image are tested consecutively against all nodes within the cascade. Whereas, a window passing all nodes is considered as a possibility for any targeted object [9].

In Viola-Jones cascade, every node is meant to possess a high (≈99.9%) detection percentage (fewer false-negatives) at the value of a low frequency of true-positives (≈50%). It suggests a higher degree of false-positives. At any node, a call of refusal stops identifying, while the detection method carries on if any short-lived recognition is found [7, 8]. So, the time for computation is saved hugely due to the extraordinarily lowered specified object sub-windows than unspecified object sub-windows of a picture. Therefore, stress ought to be arranged on sequentially weak classifiers until they maintain a sufficiently low percentage of false-negatives (incorrect refusal for objects of interest), assuring that just about all sub-windows consisting of targeted objects pass all nodes of cascaded classifiers [9].

Our goal of this research is to build a model that can detect and label objects utilizing cascade training and detection. The training is to complete using Cascade Trainer Graphical User Interface (GUI), and the testing by MATLAB. The model is translated then for implementation on the S32K234 Evaluation Board (EVB), which is finally tested for the proof of concept.

The rest of this manuscript is ordered as stated; unit II provides some connected work as well as the premise for HAAR-type options. The associate object detection algorithmic program supported the primitive cascaded classifier is projected very well in unit III. The outcomes are mentioned in unit IV, and at last in unit V, conclusions are extracted.

II. BACKGROUND

There are many kinds of algorithms deployed to identify objects or human faces with expressions as well [1, 6, 9]. All of the methods have their benefits and drawbacks [10-19]. In 2002, T. Ojala & others used Local Binary Pattern (LBP) to classify the textures of images [10]. The LBP is a grayscale insignificant texture driver with effectual distinction, additionally has been employed for the recognition of human faces. Although LBP options have excellent distinctive efficacy, they dropped the native composition underneath particular conditions.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Some strategies concerned the templates, filter, and neural network [11-13], such as the analysis that was administered by Thai Hoang Le [11]. The downside of this applied rule is that the system is too costly. Moreover, this process needs re-analyzation of the pixels of pictures for the measure and efficiency that rely upon the worth of color and strength of light, requiring an extended time, which is usually exhausting to achieve. Since 2012, several investigations occurred victimizing identical practices [14-19], one amongst them is the research by W. Sulistiyo & others that created a technique for reporting participation [16]. And, the hardware is simply a model used for attendance, except as the guard of the entrance. Furthermore, human face recognition with hijab and spectacles has not been accounted for therein analysis.

Indicators supporting the HAAR-like methodology have a considerably low rate for failing detections. The procedure operates properly and identifies objects, revolving at an angle under 30° [4]. Once the angle crosses 30°, the value of strong detections slumps. So, utilizing the quality function of this approach, it is unlikely to identify the object of interest, revolved at any angle. This specificity offers the appliance of this method sophisticated for contemporary industrial technologies [5]. In this experimentation, we presented an object detection and labeling formula in regards to a smart kitchen.

III. Model Implementation

The prototype was executed by integrating the selected software-hardware containing preprocessing and training of images of the required objects. Utilizing a microcontroller is incredibly useful because it can be programmed effortlessly. Individuals with the minimum programming experience will be able to code it comfortably because it includes all the necessary options [20-22].

A. Software & Hardware

1) Cascade Trainer

Cascade Trainer GUI is an algorithm accustomed to train, test, and develop cascade classifier designs. It generally utilizes a graphical interface to fix the values and formulate it to be simple to apply OpenCV mechanisms to train and test classifiers. The cascade classifier is a branch-formatted method, within which Viola-Jones used HAAR-type selections for identification.

The usual HAAR-type selections established by standards, as illustrated in Fig. 1, maybe deployed considering every measure within the boosted classifier for quick computation from the associated integral version of the targeted picture for detection. HAAR-type selections are deep-seated by the HAAR wavelet. A HAAR wavelet is a unique quadrilateral formation (one with large interim and one with small interim) [23].

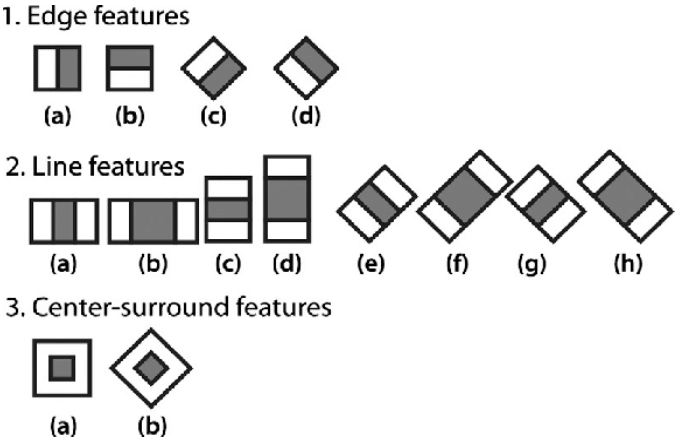


Fig. 1. HAAR-type selections wavelet illustration from OpenCV Cascade Classification (light section indicates 'add that region' and dark section indicates 'subtract that region') [23].

In 2 Dimension (D), the rectangle is represented by white and black. The significance of HAAR-type selections endures if the outcome is more than the outset, by decreasing the pixel's mean rate in black and white squares. These traits might be discovered in any space among a scanning picture window.

2) MathWorks MATLAB

The object identifying algorithms typically employ CV techniques to find and classify objects from photographs. By utilizing MATLAB (R2018b), the most suitable object identification procedures with minimal lines of code can be explored. Besides, it will be able to stimulate the method of training by the usage of GPUs [24].

Once the Cascade Trainer generates the .xml file after completing the training, MATLAB reads the file to detect cups with saucers. The vision.CascadeObjectDetector() API is used to create the detector object, which is further fed into the step() method to run the detector object to detect the object of interest. After the detection, the detected region is outlined with a rectangle and captioned as 'cup' to be displayed.

3) S32V234 EVB

The S32V234-EVB2 is a development board with maximum execution, enclosed perception, and detector fusion appliances. Supported by 64-bit Arm Cortex-A53 S32V processors, this EVB allows a pair of board choices (including/excluding integral monitor); satisfies all options, together with Automotive LAN, VIU camera, FlexRay, etc. Available are expansion card choices to assist the camera detectors also, as illustrated in Fig. 2, whereas, by deploying the cost-competitive with small form factor EVB, a vision-based case can be solved by its advantages [25].

Fig. 2. S32V234-EVB2 Evaluation System from NXP Semiconductors [25].

B. Dataset & Preprocess

A model was trained to detect and label cups, specifically – cups with saucers with different colors from other objects, as illustrated in Fig. 3. For this purpose, a set of 50 positive images were taken to train the model, collected from Google Images. Another set of images was used in the 'cup' category from Caltech101 as well. The google images yielded higher efficiency in detecting and labeling the cups. Thus, they were utilized for detecting and labeling the object of interest – cups with saucer in this case. Besides, for training purposes, 100 negative images were employed, whereas for preprocessing, all the collected low-noise images were resized to 300×300 pixels.

Fig. 3. Cup with saucer.

C. Setting & Training

For training purposes, the Cascade Trainer GUI (version 3.3.1) was used, as illustrated in Fig. 4.

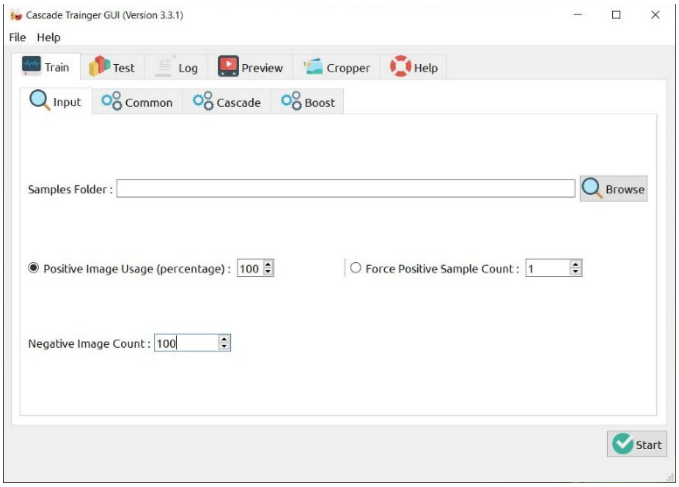


Fig. 4. The Cascade Trainer GUI interface.

The directory to the positive images was placed on the 'Sample Folder:' field of the interface. After this directory was set, the 'Negative Image Count:' field was placed to 100, as 100 negative images were taken. Once all the input specifications were set, the next step was to place the number of training stages, under the 'Common' tab, as illustrated in Fig. 5.

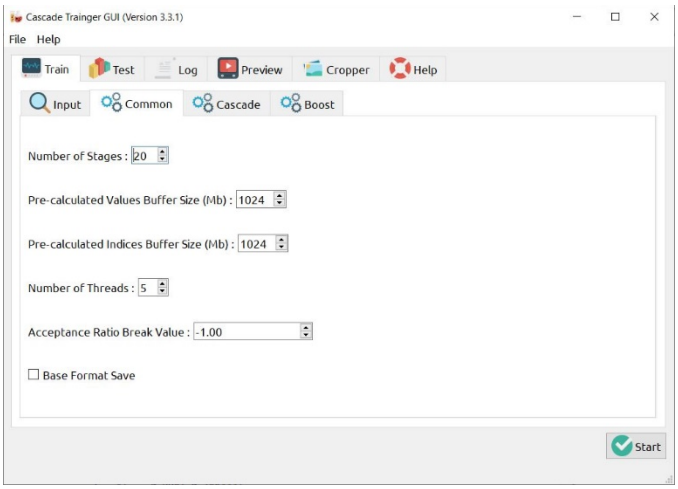


Fig. 5. Common tab specifications.

The 'Number of Stages:' was set to 20 – the higher is the number, the better is the detection accuracy. However, it takes an increasingly long time with every increment. Here, the next step was to place the parameters under the 'Cascade' tab, as illustrated in Fig. 6.

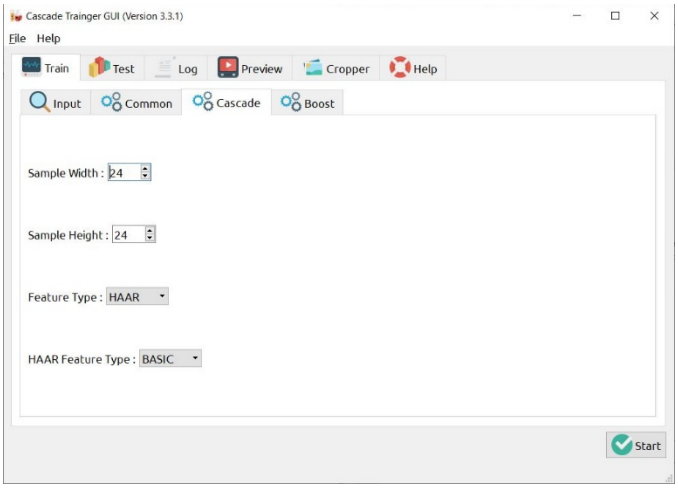


Fig. 6. Cascade training parameters.

The 'Sample Width:' and 'Sample Height:' represents the width and height of the training samples in pixels, respectively. For this model, the 'Sample Width:' and 'Sample Height:' were set to 30. The 'Feature Type:' was placed to 'HAAR,' and the 'HAAR Feature Type:' was set to 'BASIC'.

Once the training parameters were placed, the 'Start' button was deployed to start the training, which usually leads the software to use 'opencv_createsamples' internally to create positive samples. Then it employs an 'opencv_traincascade' application for training the model. When the training completes,

it gives out a .xml Cascade Classifier file for detecting cups in images through MATLAB.

IV. Experimental Result

The trained model was tested in both MATLAB R2018b and S32V234 EVB, where the processor, RAM, and OS of the deployed computer were Intel Core i5 8th Generation, 8GB, and Windows 10 Home, respectively. Both the systems successfully detected the cups with varying accuracy. Fig. 7 shows the MATLAB output of the tested images. It can be seen that the model detected all the cups when running on MATLAB with perfect accuracy. However, Fig. 8 shows a reduction in the accuracy of detection in terms of S32V234 yield.

Fig. 7. The MATLAB output.

Fig. 8. The S32V234 output.

To test the developed system, random pictures were used, such as a tree, human brain, and cups with saucers with separate colors. The detection approach was justified for a wide range of possibilities by the usage of a set of mixed objects. Fig. 9 shows an instance of a false-positive output generated when the model was run on the EVB.

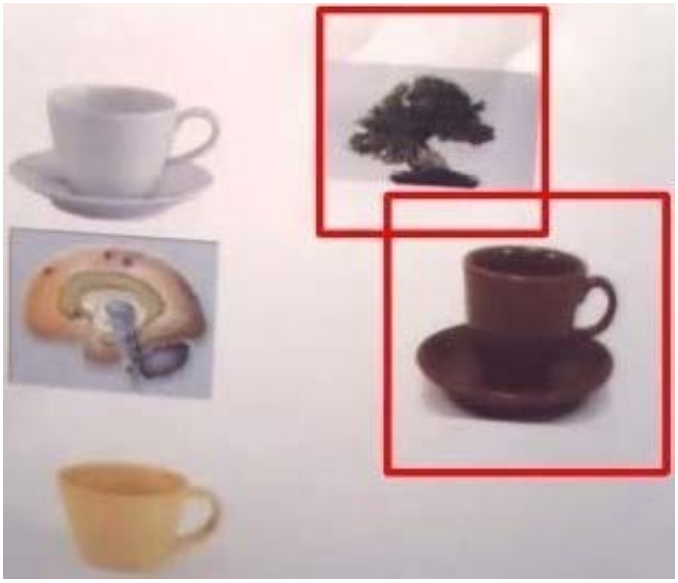


Fig. 9. S32V234 false-positive outcome.

When the Caltech101 images were used for training the model, the detection did not yield a satisfactory result. Upon closer inspection, it was identified that the Caltech101 images were noisy, and the edges, as well as the corners, were not distinct enough to be detected by the Cascade Trainer GUI. After training, the model was not able to detect any of the positive images employed for training; thereby, the Google Images results were utilized, which produced better outcomes in the detection of the object of interest.

In the case of testing, the model showed different results for MATLAB and S32V234. The discrepancies in the results could be caused by improper preprocessing of the images. There was an instant when all the cups were detected by the board, but the detection stopped before the screenshot was taken. This might be indicating the lack of capability to operate in real-time for the trained model.

V. Conclusion

The research was done to train a model to detect some objects of interest among diversified objects, which in this case, were cups with saucers with differing colors. Our developed algorithm in terms of CV successfully detected and labeled the test images with varying degrees of accuracy – neutralizing the deficiencies of the primordial Viola-Jones' cascade classifier while maintaining almost zero incorrect refusal; however, there are scopes for improvement. The principles utilized for recognizing the cups could be employed for detecting other objects as well, in a multitude of scenarios.

To enhance the accuracy of the model, more extensive training with a large dataset should be carried out – increasing

the number of positive and negative images. Higher computing power could be used, and the discussed number of stages along with samples might be increased to upgrade the learning rate. A comparative study would be carried out by utilizing other training methods for achieving higher performance regarding object detection.

ACKNOWLEDGEMENT

We would like to extend our humble gratitude towards Dr. Mohamed El-Sharkawy – Professor of Electrical & Computer Engineering, Purdue School of Engineering & Technology, for his resolute pedagogical support. Without his support, the compiled work would not have been possible. Furthermore, we also offer our sincerest appreciation towards the Internet of Things Collaboratory lab for all the hardware support which has proven paramount for all the testing and proof of concepts.

REFERENCES

[1] H. Schneiderman, "Feature-centric evaluation for efficient cascaded object detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004, vol. 2, doi: 10.1109/cvpr.2004.1315141.

[2] E. T. Efaz, A. Al Mamun, K. Salman, F. Kabir, S. N. Sakib, and I. Khan, "Design of an indicative featured and speed controlled obstacle avoiding robot," in *2019 International Conference on Sustainable Technologies for Industry 4.0, STI 2019*, Dec. 2019, doi: 10.1109/STI47673.2019.9068018.

[3] E. T. Efaz, M. M. Mowlee, J. Jabin, I. Khan, and M. R. Islam, "Modeling of a high-speed and cost-effective FPV quadcopter for surveillance," in *23rd International Conference on Computer & Information Technology, ICACIT 2020*, Dec. 2020.

[4] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, vol. 1, doi: 10.1109/cvpr.2001.990517.

[5] P. Viola and M. Jones, "Robust real-time face detection," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2, p. 747, 2001, doi: 10.1109/ICCV.2001.937709.

[6] D. S. Chen and Z. K. Liu, "Generalized haar-like features for fast face detection," in *Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, ICMLC 2007*, 2007, vol. 4, pp. 2131–2135, doi: 10.1109/ICMLC.2007.4370496.

[7] R. Yustiawati *et al.*, "Analyzing of Different Features Using Haar Cascade Classifier," in *Proceedings of 2018 International Conference on Electrical Engineering and Computer Science, ICECOS 2018*, Jan. 2019, pp. 129–134, doi: 10.1109/ICECOS.2018.8605266.

[8] D. Peleshko and K. Soroka, "Research of usage of Haar-like features and AdaBoost algorithm in Viola-Jones method of object detection," *2013 12th International Conference on the Experience of Designing and Application of CAD Systems in Microelectronics (CADSM)*, Polyana Svalyava, 2013, pp. 284-286.

[9] C. Li, Z. Qi, N. Jia, and J. Wu, "Human face detection algorithm via Haar cascade classifier combined with three additional classifiers," in *ICEMI 2017 - Proceedings of IEEE 13th International Conference on Electronic Measurement and Instruments*, Jul. 2017, vol. 2018-January, pp. 483–487, doi: 10.1109/ICEMI.2017.8265863.

[10] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002, doi: 10.1109/TPAMI.2002.1017623.

[11] T. H. Le, "Applying Artificial Neural Networks for Face Recognition," *Adv. Artif. Neural Syst.*, vol. 2011, pp. 1–16, 2011, doi: 10.1155/2011/673016.

[12] A. S. Abstrak, "SISTEM PENGENALAN WAJAH MENGGUNAKAN METODE PRINCIPAL COMPONENT ANALYSIS (PCA) DENGAN ALGORITMA FUZZY C-MEANS (FCM)," *J. Pendidik. Mat.*, vol. 4, no. 2, pp. 58–65, 2015.

[13] Wirtjes and S. Jaceline, "Pengenalan Ekspresi Wajah Menggunakan Convolutional Neural Network (CNN)," *Repos. Institusi USU*, vol. 4, no. 3, pp. 4907–4916, 2019, [Online]. Available: http://repositori.usu.ac.id/handle/123456789/15450.

[14] R. Padilla, C. C. Filho, and M. Costa, "Evaluation of Haar Cascade Classifiers for Face Detection," *J. WASET*, vol. 6, no. 4, pp. 323–326, 2012.

[15] V. Singh, V. Shokeen, and B. Singh, "FACE DETECTION BY HAAR CASCADE CLASSIFIER WITH SIMPLE AND COMPLEX BACKGROUNDS IMAGES USING OPENCV IMPLEMENTATION," *Int. J. Adv. Technol. Eng. Sci.*, vol. 1, no. 12, pp. 33–38, 2013.

[16] W. Sulistiyo, B. Suyanto, I. Hestiningsih, - Mardiyono, and - Sukamto, "Rancang Bangun Prototipe Aplikasi Pengenalan Wajah untuk Sistem Absensi Alternatif dengan Metode Haar Like Feature dan Eigenface," *JTET (Jurnal Tek. Elektro Ter.*, vol. 3, no. 2, Aug. 2014, doi: 10.32497/JTET.V3I2.180.

[17] H. H. Lwin, A. S. Khaing, and H. M. Tun, "Automatic Door Access System Using Face Recognition," *Int. J. Sci. Technol. Res.*, vol. 4, no. 6, pp. 294–299, 2015.

[18] H. Yang and X. A. Wang, "Cascade classifier for face detection," doi: 10.1177/1748301816649073.

[19] I. Gangopadhyay, A. Chatterjee, and I. Das, "Face Detection and Expression Recognition Using Haar Cascade Classifier and Fisherface Algorithm," in *Advances in Intelligent Systems and Computing*, 2019, vol. 922, pp. 1–11, doi: 10.1007/978-981-13-6783-0_1.

[20] J. Jabin, A. M. Chowdhury, E. T. Efaz, M. E. Adnan, and M. R. Islam "An automated agricultural shading for crops with multiple controls," in *2020 International IOT, Electronics & Mechatronics Conference, IEMTRONICS 2020*, Sep. 2020.

[21] M. E. Adnan, N. M. Dastagir, J. Jabin, A. M. Chowdhury, and M. R. Islam, "A cost effective electronic braille for visually impaired individuals," in *5th IEEE Region 10 Humanitarian Technology Conference 2017, R10-HTC 2017*, Feb. 2018, vol. 2018-January, pp. 175–178, doi: 10.1109/R10-HTC.2017.8288932.

[22] J. Jabin, M. E. Adnan, S. S. Mahmud, A. M. Chowdhury, and M. R. Islam, "Low cost 3D printed prosthetic for congenital amputation using flex sensor," in *2019 5th International Conference on Advances in Electrical Engineering, ICAEE 2019*, Sep. 2019, pp. 821–825, doi: 10.1109/ICAEE48663.2019.8975415.

[23] "Cascade Classification — OpenCV 2.4.13.7 documentation." https://docs.opencv.org/2.4/modules/objdetect/doc/cascade_classification.html (accessed Aug. 06, 2020).

[24] "What is Object Detection Video - MATLAB & Simulink." https://www.mathworks.com/videos/what-is-object-detection-1564383482370.html (accessed Aug. 06, 2020).

[25] "S32V Vision and Sensor Fusion Evaluation System | NXP." https://www.nxp.com/design/development-boards/automotive-development-platforms/s32v-mpu-platforms/s32v-vision-and-sensor-fusion-evaluation-system:S32V234EVB (accessed Aug. 06, 2020).

An Adaptive Overlap-Pipelined Multitasking Superscalar Processor

Mong Tee Sim
Dept of Computer Science
University of Colorado Colorado Springs
Colorado, USA
mong_sim@hotmail.com

Qing Yi
Dept of Computer Science
University of Colorado Colorado Springs
Colorado, USA
qyi@uccs.edu

***Abstract*—Low power consumption, high performance, and small die size are the three essential considerations in modern CPU design, from tiny IoT devices to General Purpose Manycore System-on-Chip. With these considerations, we introduce a new CPU design that features Adaptive Overlapping Multitasking pipelines, to better balance the design tradeoffs of the traditional scalar and superscalar CPUs. By providing dynamic reconfigurability, we enable user applications to decide at run-time whether to run the CPU in a high-performance or a low-power mode, to meet their respective application deadlines or power budgets. The low-power mode also provides redundancies that allow the CPU to continue operating, even when some of its pipeline stages are damaged. We used the RISC-V ISA test suite, Dhrystone, Coremark, and ten other benchmarks to validate our CPU design's functionality and performance. Our CPU can consistently deliver up to 2.0 Instruction Per Cycle and score a 3.924 DMIPS/MHz and 6.556 Coremark/MHz with Dhrystone and Coremark benchmarks.**

Keywords—Superscalar, Multitasking, Fine-grained, Manycore, Multicore

I. INTRODUCTION

Low power consumption and high performance are essential for embedded applications, especially when their power source comes from the battery or solar panel. Reducing the power consumption ratio is vital to multicore and manycore CPUs, where power consumption can be a significant issue. Smaller die size, however, is also a concern in CPU design, especially in manycore systems. To measure the efficiency and performance, we use CPU performance to refer to how fast a CPU can operate in terms of clock frequency, and use CPU efficiency to refer to the number of instructions it can process per clock cycle. In this context, CPU design has followed two directions: the Speed Demon [5] (super pipelining) approach and the Brainiac [5] (superscalar) approach. The Speed Demon approach subdivides its CPU stages to reduce the resultant stage latency and boost the clock frequency. It, therefore, increases the number of processed instructions per time unit. The Brainiac approach, which issues multiple instructions per cycle without increasing the clock frequency, improves the average instruction per cycle (IPC). However, both methods consume more power due to higher operating frequency and more hardware.

Since the Speed Demon CPUs reduce their stage latencies to execute at a higher clock frequency, the power consumption rate increases proportionally. For example, the Intel Prescott family processors peaked up to 31 pipeline stages, consumes up to 213W, and clocked at a 3.60 GHz. When the pipeline stages stall, e.g., due to data dependencies, the system's throughput decreases. With the addition of more pipeline stages and new functional blocks, the higher operating frequency also increases the overall power consumption.

On the other hand, the Brainiac (superscalar [11] [12] [13][14]) approach issues multiple instructions per clock cycle to enhance CPU throughput without increasing the clock frequency. The micro-architecture design includes various instruction issuing policies and registers renaming algorithms to improve CPU throughput. However, its wide instruction fetching unit also requires a complex instruction cache and branch prediction design to reduce clock cycles. Since the number of independent instructions is often not substantiated, the maximum CPU throughput is often not realized.

Our research focuses on building a quad-core CPU targeting embedded and cyber-physical applications and manycore systems. With die size reduction for manycore systems. We design a dual-issue overlapping superscalar processor operating in fine-grained multitasking [2][6][7][8] fashion to improve its efficiency, and trade higher frequency for higher IPS (Instruction per Second) and smaller die size. Our new dual-issue superscalar CPU leverages a fine-grained multitasking architecture to eliminate the need for a branch predictor (BP) [9] and Return Address Stack (RAS) [10][15], allowing it to reduce the overall power consumption and improve system efficiency.

Our new microarchitecture supports fine-grained multitasking, similar to a classic five-stage pipeline, in terms of available hardware and operations. In particular, we loosely couple the two pipelines to allow run-time reconfiguration that makes the CPU work either in a coupling high-performance mode or a decoupling low-power mode. Our new design provides a redundancy option in the decoupling mode that will keep the CPU running if one of the pipeline stages is damaged. In the coupling mode, our new design operates as a dual-issue multitasking superscalar processor.

Since the available number of independent instructions is often unknown in superscalar mode, our CPU fetches from two distinct instruction streams simultaneously when the CPU runs more than two virtual-cores. This instruction issuing paradigm eliminates our dual-issue superscalar processor's true-data dependency while creating virtualization of overlapping

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

pipelines. Our dual-issue multitasking CPU design provides dynamic reconfiguration capability that enables the user application to decide at run-time for the CPU to operate in its low-power or high-performance mode, to meet varying deadlines or power budgets.

Our research makes the following contributions:

- Our new design can achieve an IPC up to 2.0, reaching the theoretical limit of a dual-issue superscalar processor.
- We use a dual-core fine-grained architecture to create a quad-core architecture and trade higher operating frequency for higher IPS, and smaller die size. This technique allows us to reduce the die size by more than 50%, making it ideal for manycore implementation.
- The CPU decoupling feature provides a redundancy capability that allows the CPU to continue to operate even when some of its pipeline stages are damaged.
- Our design requires less than half the size of a typical quad-core. Even at a higher operating frequency, its power consumption is lower than the other quad-core designs.

The rest of this paper is organized as follows: Section II briefly introduces how fine-grained multitasking works. Section III summarizes the new micro-architecture design. Section IV elaborates on our new design operating in different modes. Section V evaluates our CPU design's efficiency by using the Dhrystone, Coremark, and ten other benchmarks, Section VI summarizes some related work. Section VII concludes with some discussions.

II. The Fine-grained Multitasking Architecture

Our fine-grained multitasking architecture (FGMA) enables a single CPU pipeline assembly to support multiple real-time threads or virtual cores (VC), each VC isolated by the hardware and shielded from the other VCs [2]. If any VC malfunctions, e.g., due to errors in the application code, the other active VCs would not be affected and can continue to operate as usual. Therefore, user applications can treat the VCs as hardware isolated real-time threads that execute in an interleaved fashion, as a virtual multicore system that supports multiple standalone applications, or a mix of both.

To support fine-grained multitasking, our CPU interleaves all active VCs (hardware-isolated threads), fetching one instruction each time for a different VC in a round-robin fashion. The interleaving paradigm provides hardware support for context switching among the VCs, without requiring any runtime overhead. This paradigm also introduces stage delays (see Fig. 1) among the running VCs to mitigate any load-used data hazards (due to memory loads) and any control hazards (due to branching). As a result of these stage delays, there are no need for hardware-supported BP [9] or RAS [10] [15] modules (properties of FGMT), reducing power consumption and die size of our system.

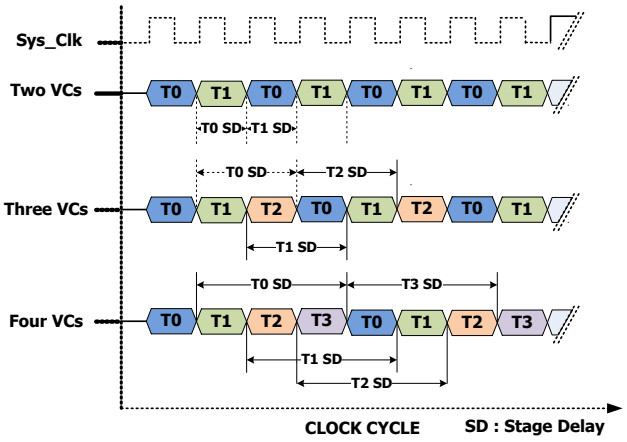


Fig. 1: This diagram shows how the threads (TX) execution in our new design. In this paradigm of executing threads, it also produces stage delays. The stage delay of a thread is defined as when the system is running other threads (or virtual cores).

Fig. 1 illustrates how our CPU evaluates the instructions when equipped with two pipelines. To support additional fine-grained multitasking, e.g., for up to four virtual cores, we need to augment the CPU with four program counters, four sets of register files, and a two-bit virtual core identification (VCID). We use switching fabrics that embody the four program counters in a program counter block (PCB) to operate the pipelines in an interleaved fashion. We further incorporate VCID awareness in the forwarding and hazard logic to ensure proper operation (see Fig. 3).

III. New CPU Design

Leveraging the fine-grained multi-tasking design, our CPU (see Fig. 3) includes two loosely coupled pipeline assemblies that support run-time reconfiguration. When the user application requires low power consumption, it can command the PCB to decouple the two pipelines and power down the idle one to reduce power consumption. Conversely, if the user applications require high-performance, it can issue a coupling command to the PCB to couple the two pipelines to form a dual-issue superscalar processor. The following section explains these two operating modes in more detail.

A. Low-power or Decoupled Mode

Upon power-on-reset (POR), the system enters its default single-threaded mode, enabling the top-pipeline and disabling the bottom-pipeline (see Fig. 3). In this mode, a user application can issue a runtime command to trigger the CPU to enter a multi-threaded low-power mode. The PCB provides five bits for the mode changing operations. The C-bit or the coupling bit couples/decouples the pipelines when set/cleared. The S-bit disables or enables the top-/bottom-pipelines when set/cleared. The P-bit or the power bit turns on the top-/bottom-pipelines based on the S-bit. When the S-bit toggles, the P-bit is set, powering up the other pipeline while the system disables the active pipeline. The T-bit and B-bit are read-only bits that provide the top and bottom-pipelines power status.

In low-power mode, only one pipeline can execute instructions, and the other is powered down. The user application must decide which pipeline should stay active and

which to power down by clearing/setting the S-bit, followed by de-asserting the C-bit. For example, assume the CPU is currently in the coupling (high performance) mode. To have the top-pipeline continue to stay active and the bottom-pipeline powered down, the user application must first clear the S-bit (selecting top-pipeline), followed by clearing the C-bit (entering the decoupled mode). The PCB then de-asserts a "disable-bottom" control signal and polls the T and B status bits. The "disable-bottom" control signal is then propagated through the pipelined stages, each of which is disabled under the control of the hazard control logic. The control hazard ensures each instruction completes before the "disable-bottom" control signal disables its pipelined stage. The "disable control" signal is fed back to the PCB at the WB stage as the B-bit. Once the T- and B-bits values are 0b10, the PCB starts issuing instructions only to the top-pipeline while disabling the bottom-pipeline.

When switching from the top-pipeline to the bottom one, a user application needs to issue a command to change the pipeline by setting the S-bit in the PCB. Upon issuing this command, the PCB sets the P-bit to power up the bottom-pipeline, stops issuing instructions to the CPU, de-asserts the "disable-top" control signal, and waits for feedback from the WB stage. Once the T- and B-bits values are 0b01, the PCB resets the bottom-pipeline, starts issuing instructions to the bottom-pipeline, and powers down the top-pipeline.

The pipeline switching feature allows user applications to continue to run on a different pipeline even if any stages in one pipeline is damaged. This pipeline validation and switching task should be one of the Designs for Test (DFT) tasks. In this scenario, the system cannot achieve the superscalar mode (coupling mode).

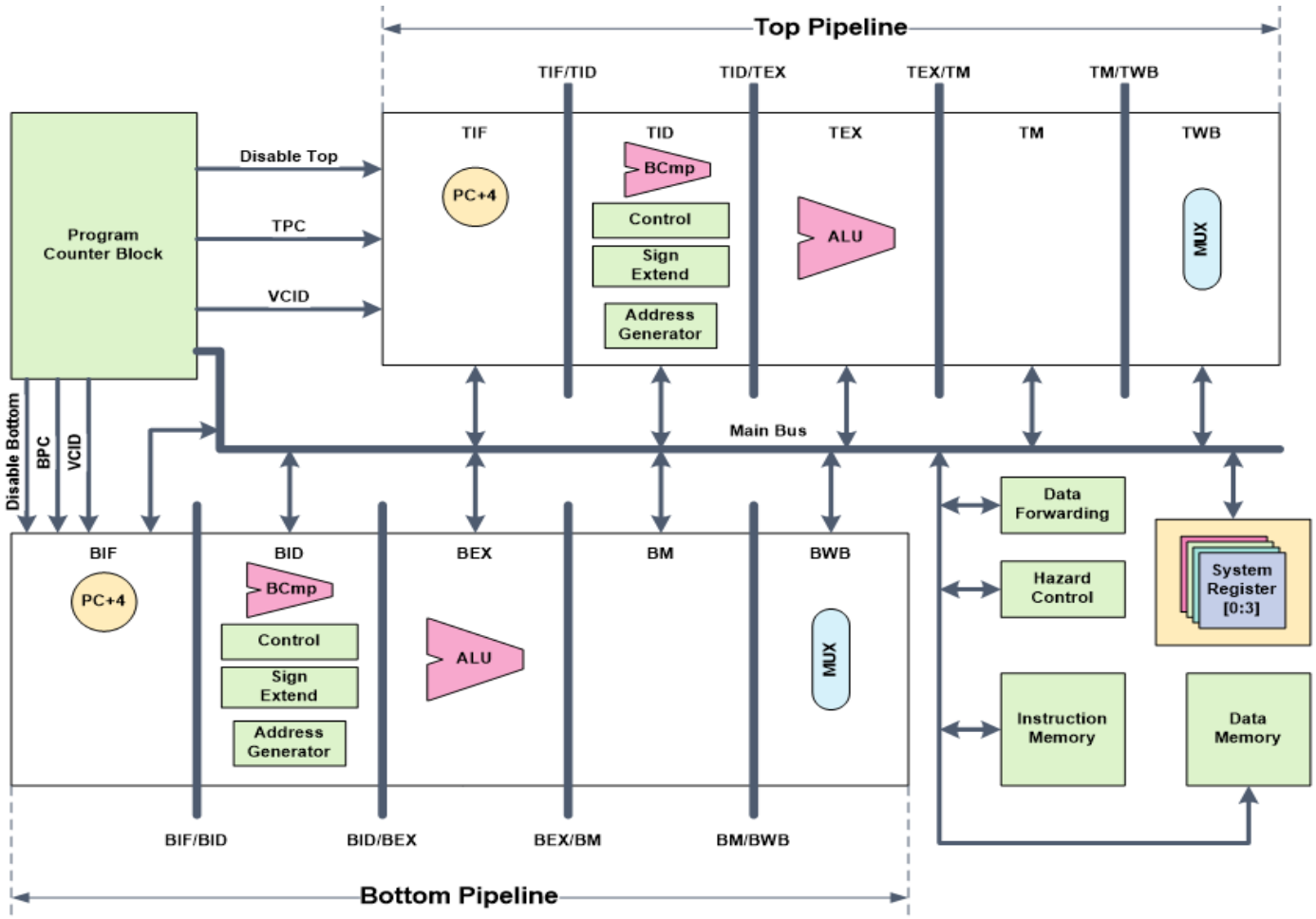


Fig. 2: This block diagram shows our Adaptive Overlapping Multitasking Superscalar Processor. The Program Counter Block consists of the switch fabric and four Program Counters. There are four system register files, each to support a running thread or virtual core.

B. High-performance or Coupling Mode

In the high-performance mode, the system couples the two pipelines to form a dual-issue superscalar processor. When a user application requires high performance, it issues a coupling command to the PCB. Upon receiving the coupling command, the PCB powers up the idle pipeline, and waits for the pipeline WB stage feedback. Once the T- and B-bits values are 0b11, the PCB resets the recently powered up pipeline and then starts issuing instructions to the superscalar processor.

IV. Operating in Superscalar Mode with Two to Four VCs

This section demonstrates how our new design achieves an average IPC of 2, reaching the theoretical limit of a dual-issue superscalar processor.

Upon POR, the CPU enters the low-power single-threaded mode. In this state, the system behaves like a five-stage pipelined scalar processor without a BP (operating in branch always not taken) and RAS support. The system runs on VC 0, and its related resources such as PC, register file (see Fig. 2). In this state, the user application can configure the system as a multi-core system running multiple standalone applications, or a multitasking system running multiple hard real-time threads, or a mix of both (see Fig. 1).

The superscalar mode (i.e., the fine-grained multitasking mode) can support up to four VCs. In this mode, the PCB uses different instruction issuing algorithms to control the PCs, and with the address forwarding hardware when running more than two VCs. We will demonstrate how the instruction issuing algorithms work for two to four VCs in this section.

A. *Supporting Two Virtual Cores*

CC	TIF	TID	TEX	TM	TWB	
0	VC0					T
	VC0					B
1	VC1	VC0				T
	VC1	VC0				B
2	VC0	VC1	VC0			T
	VC0	VC1	VC0			B
3	VC1	VC0	VC1	VC0		T
	VC1	VC0	VC1	VC0		B
4	VC0	VC1	VC0	VC1	VC0	T
	VC0	VC1	VC0	VC1	VC0	B
5	VC1	VC0	BT	VC0	VC1	T
	VC1	VC0	BS	VC0	VC1	B
	BIF	BID	BEX	BM	BWB	

Fig. 3: When in the Superscalar mode running two VCs, the system issues two instructions to the CPU pipelines. The pipelines are arranged logically in parallel (CC: Clock Cycle, BS: Bubble Stage, BT: Branch Taken, T: Top-pipeline, and B: Bottom-pipeline).

Fig. 3 shows our new design supports two VCs in the FGM superscalar mode. The system issues two instructions from both applications in an interleaved fashion and repeats in an infinite loop. From Fig. 3, from CC=0 to CC=4, we assume that the system is hazard-free and running VC0 and VC1. This method of issuing instructions inherits the true-data dependency issue that cannot be resolved by data forwarding. The system must stall and wait for the respective stage to produce the needed result. Instead of stalling, we evict the true-data dependency instruction at the ID stage to minimize the efficiency degradation and implementation complexity.

When CC=4 (see Fig. 3), the TID (the first instruction in the top-pipeline) decodes the branch taken, and evicts the second instruction in the bottom-pipeline BID stage, by inserting a bubble stage on the next cycle (the BS stage when CC=5) at the BEX stage. The branch-taken (BT) instruction on the top-pipeline has no function after the TID stage. Overall, the system efficiency of our CPU can be reduced in the following scenarios:

- If the first instruction the system decodes is a jump (including return) or a branch taken, the system executes the first instruction and discards the second instruction.
- If the second instruction the system decodes depends on the first instruction result, the system executes the first instruction and discards the second instruction.

B. *Supporting Three Virtual Cores*

CC			TIF	TID	TEX	TM	TWB	
0								T
	VC0							B
1								T
	VC1	VC0						B
2			VC0					T
	VC2	VC1	VC0					B
3			VC1	VC0	VC2	VC1		T
	VC0	VC2	VC1	VC0				B
4			VC2	VC1	VC0	VC2	VC1	T
	VC1	VC0	VC2	VC1	VC0			B
5			VC0	VC2	VC1	VC0	VC2	T
	VC2	VC1	VC0	VC2	VC1			B
6			VC1	VC0	VC2	VC1	VC0	T
	VC0	VC2	VC1	VC0	VC2			B
	BIF	BID	BEX	BM	BWB			

Fig. 4: In the superscalar mode running three VCs, the system issues the first instruction to the bottom-pipeline. The second instruction from the same program is fetched based on the prior instruction outcome in the BID stage (T: Top-pipeline, and B: Bottom-pipeline).

We illustrate how the PCB issues instructions to both pipelines after POR in Fig. 5, which shows our new design is running three VCs. In this mode, our system has one more stage delay of freedom (see Fig. 1, three VCs, two-stage delays). The PCB takes advantage of this additional stage delay to issue instructions from two programs, one in each clock cycle. The PCB issues the first instruction to the bottom-pipeline, and the address to fetch in the next instruction depending on the outcome of the first instruction when it reaches the BID stage (Fig. 5, when CC=1 and CC=2).

Since the two instructions from the same program are two stages apart, we can resolve data dependencies with data forwarding; hence, this method mitigates true-data dependency. Fig. 6 shows that when the first instruction is at the BID stage (bottom-pipeline ID stage), the instruction decoder decodes the instruction and makes the following decision:

- If the first instruction the BID decodes is not a branch or jump instruction (including return), the BID will provide the value of PC + 4 to TIF to fetch as the next instruction.
- If the first instruction decoded by the BID is a branch or jump instruction, the BID will provide the branch/jump target address to the TIF to fetch as the next instruction.

This method of utilizing the BID to provide the address to TIF also improves the IPC of this configuration. However, this configuration suffers from performance degradation in two scenarios. First, if a branch instruction in the BID depends on the previous TEX's instruction result (Fig. 5, when CC=X), then the system will stall. Instead of stalling, the BID predicts a branch-taken and provides the TIF target address to fetch the

next instruction (Fig. 5, when CC=X + 1). If the prediction is wrong (branch-not-taken), the system flushes the TIF and BIF instructions (Fig. 5, CCX + 2). The BIF instruction gets into the bottom-pipeline before the relative PC can correct its address (from the branch-not-taken path, shown in Fig. 5, when CC=X + 2). The BNT path (Fig. 5, when CC=X + 2) has no function after the BID stage.

CC			TIF	TID	TEX	TM	TWB	
X			VC2	VC1	VC0	VC2	VC1	T
	VC1	VC0	VC2	VC1	VC0			B
X+1			VC0	VC2	VC1	VC0	VC2	T
	VC2	VC1	VC0	VC2	VC1			B
X+2			VC1	BS	VC2	VC1	VC0	T
	BS	VC2	VC1	BNT	VC2			B
	BIF	BID	BEX	BM	BWB			

Fig. 5: The wrong prediction occurs when Branch-not-taken (BNT). The system evicts the residual instruction by inserting a bubble stage (BS: Bubble stage, T: Top-pipeline, B: Bottom-pipeline)

CC			TIF	TID	TEX	TM	TWB	
Y			VC1	VC0	VC2	VC1	VC0	T
	VC0	VC2	VC1	VC0	VC2			B
Y+1			VC2	VC1	BT	VC2	VC1	T
	VC1	BS	VC2	VC1	VC0			B
	BIF	BID	BEX	BM	BWB			

Fig. 6: The wrong prediction occurs when branch was taken. The system evicts the residual instruction by inserting a bubble stage (BS)

The second performance degradation scenario happens when the next instruction on the top-pipeline is a branch-taken instruction. At the TID stage (Fig. 6, when CC=Y), the TID decodes a branch-taken. In the same clock cycle, the BIF fetches its next instruction for the bottom-pipeline, assuming a normal operation. Under this condition, the system evicts the residual BIF instruction (Fig. 6, when CC=Y) from the bottom-pipeline in the next clock cycle by inserting a bubble stage in the BID (Fig. 6, when CC=Y + 1). In this case, the system efficiency degrades by one instruction.

C. Supporting Four Virtual Cores

Fig. 7 shows when our new design is running four VCs. The issuing paradigm in this configuration is similar to when running three VCs, except that it has three-stage delays between instructions. The additional stage delay helps mitigate all the hazards experienced by the three VCs configuration.

In Fig. 7, when CC=5, the BID (VC0) decodes a branch instruction dependent on an earlier instruction's result. The system can forward the result from the TM (Fig. 7, when CC=5) to the BID to complete its decoding. In this case, we resolve the wrong prediction issue in the three VCs mode.

CC			TIF	TID	TEX	TM	TWB	
0								T
	VC0							B
1								T
	VC1	VC0						B
2			VC0					T
	VC2	VC1	VC0					B
3			VC1	VC0				T
	VC3	VC2	VC1	VC0				B
4			VC2	VC1	VC0			T
	VC0	VC3	VC2	VC1	VC0			B
5			VC3	VC2	VC1	VC0		T
	VC1	VC0	VC3	VC2	VC1			B
6			VC0	VC3	VC2	VC1	VC0	T
	VC2	VC1	VC0	VC3	VC2			B
7			VC1	VC0	VC3	VC2	VC1	T
	VC3	VC2	VC1	VC0	VC3			B
	BIF	BID	BEX	BM	BWB			

Fig. 7: In the Superscalar mode executing four VCs, the system issues one instruction to the CPU bottom-pipeline and next instruction based on the BID to the top-pipeline. The pipelines logically overlap.

In this paragraph, we show that we can mitigate the three VC's second scenario by running four virtual cores (or threads). When VC3 at the TEX stage decodes the second branch (branch-taken) instruction, it updates the VC3 program counter with the branch target address. In this clock cycle, the BIF loads VC2's instruction (not related to VC3). In the next clock cycle, VC3 uses the TEX's branch target address (when CC=5) to fetch the instruction into the bottom-pipeline. This sequence of steps shows that the four VC modes also mitigate performance degradation (IV.B) in the second scenario of the three VCs configuration.

V. ANALYSIS

Our design focuses on die size reduction without compromising the performance. We trade smaller die size with higher operating frequency without increasing the power consumption for the equal number of cores for a typical multicore system. We configure our design as Quad (running four VCs) for a manycore system, to reduce its die size by half with a given die size (two cores versus four cores). In this section, we first evaluate the system performance, and then operating frequency and power consumption.

A. Benchmark Results

We used industry-standard benchmarks to validate the effectiveness of our design. In this experiment, we use the Dhrystone (Version 2.1) and Coremark (2009) benchmarking software and the riscv64-unknown-elf-gcc (Version 7.1.1) compiler. The CPU functionality is verified using the RISCV ISA test suite. Additionally, we use ten other selected benchmarks to show how our design performs when executing different algorithms. These ten benchmarks were chosen mainly due to their algorithmic nature and development in real embedded applications. These include VVADD (adds to vector algorithm), TOWER (the tower of Hanoi), SHA-256, CRC32, QSORT, MEDIAN, AES, BLOWFISH, MD5, and DES.

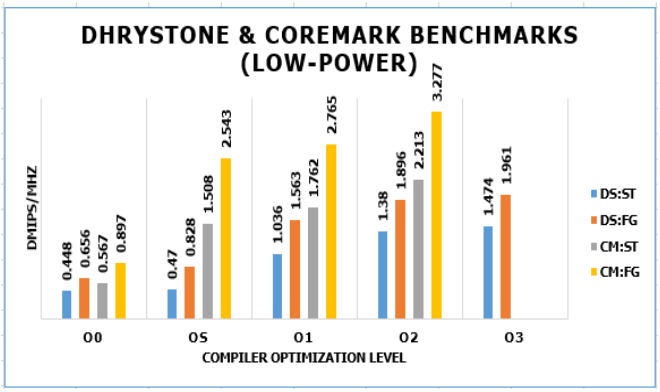


Fig. 8: Dhrystone and Coremark benchmarks results of our new design operating in low-power mode. (DS: Dhrystone and CM: Coremark)

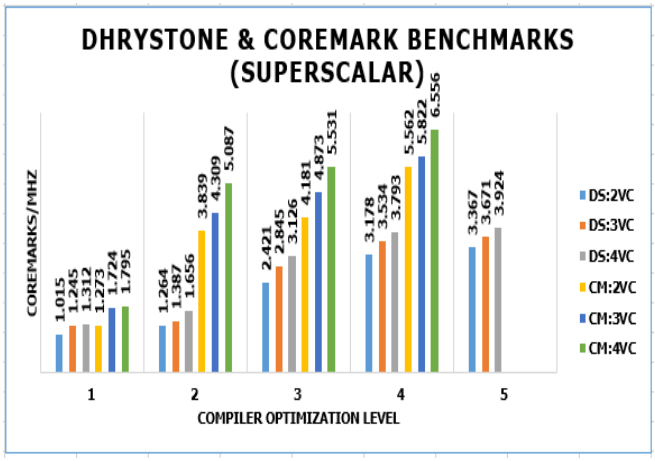


Fig. 9: Dhrystone and Coremark benchmarks results of our new design operating in superscalar mode. (DS: Dhrystone and CM: Coremark)

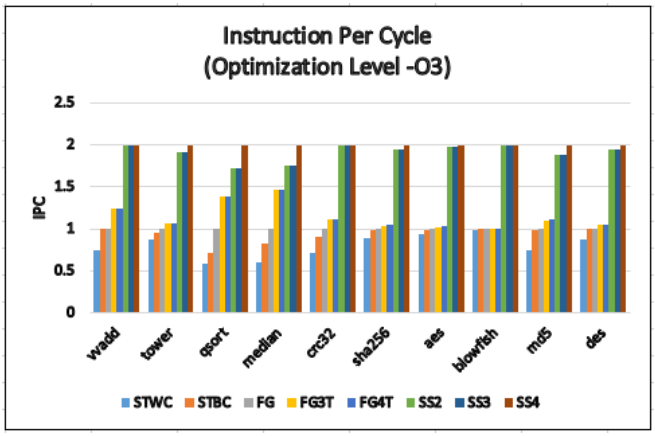


Fig. 10: Our new design IPC performance with low-power and superscalar modes and STWC and STBC (LP-Low-power or decoupled mode, SS: Superscalar or coupled mode, STWC: Scalar processor worst case, and STBC: Scalar processor best case).

Fig. 8 and Fig. 9 show the Dhrystone and Coremark benchmark scores running at different compiler optimization levels, in the low-power and superscalar modes. Our design operating in the superscalar mode can achieve the Dhrystone, and Coremark scores up to 3.924 DMIPS/MHz, and 6.556 Coremark/MHz, respectively.

The fine-grained processor does not have control and data hazards, and in this test case, there is no structural hazard. The IPC scores in Fig. 10 were obtained by dividing the number of clock cycles required to run our design in coupled and decoupled mode, by the number of cycles needed to run the equivalent benchmark on a single cycle CPU.

TABLE I: COREMARK PERFORMANCE OF OUR NEW DESIGN VERSUS THE AMSP, SERIES 7 RISC-V AND ARM PROCESSORS

	Our Design	AMSP	SiFive E76	ARM Cortex-M7	SiFive U74	ARM Cortex-A55
Instruction Set	***32-bit RISC-V***	32-bit RISC-V	32-bit RISC-V	32-bit ARM V7-M	64-bit RISC-V	64-bit ARM V8
Max IPC	***2***	2	2	2	2	2
Avg. IPC	***2.0***	1.697	1,495	1.525	1.495	1.343
CoreMark Perf CM/MHz	***6.556***	5.562	4.9	5.0	4.9	4.4

From our ten benchmarks (see Fig. 10), the STWC (single thread worst case) five-stage scalar processor without the BP and RAS has the worst performance with QSORT and MEDIAN. These two benchmarks have more control instructions relative to the other eight benchmarks. All the benchmark results confirm that our design running 4VCs' excels in performance, delivering an average IPC=2, reaching the theoretical limit of a dual-issue superscalar processor.

We compare the fine-grained and commercial superscalar processors from Sim and Yi [1], SiFive [3], and ARM [4] performance with ours. This comparison is relevant because they are all dual-issue superscalar processors. In terms of micro-architecture and the instruction issuing paradigm, our CPU outperforms Sim and Yi processor in IPC by 0.994 (6.556 – 5.562). The SiFive processors used the same ISA as in our design. We choose ARM processors because it is the current industry de facto standard with millions of sockets won.

We obtain the performance metrics from the respective researchers and companies' literature. Our design outperforms these three families of processors by a large margin (see Table 1). Based on these results, our design is also suitable for the manycore system because its die size is less than half of a traditional quad-core CPU. In the next section, we examine the operating frequency and power consumption requirement of our design to match the performance of a traditional quad-core CPU.

VI. RELATED WORK

Sim and Yi [1] provided a detailed description of how their fine-grained multitasking superscalar processor works and the different algorithms they used to improve the system efficiency. Their idea departs from the traditional superscalar, focusing on hard real-time multitasking and reducing hardware to improve both the system performance and die size.

Sim and Perera [2] provided a comprehensive description of how the fine-grained multitasking works. Their paper utilized a multitasking subsystem with a sequencer to provide both spatial and temporal multitasking. The dynamic CPU clock cycle allocation extends to user application demands allowing further optimization of the CPU clock cycle usage. The stack overflow and underflow, and the global memory protection are great features for RTOS since most RTOS does not have memory protection between threads.

Most of the other existing work [6] [7] [8] focuses on the hard real-time and the scheduling methods for fine-grained processors.

VII. Discussion and Conclusion

This paper demonstrates that our new design performs significantly better than other superscalar processors. With an overlapping pipeline and branch address forwarding logic, our processor design mitigates the true-data dependency issue. When operating in the 4VC mode, our system achieves the theoretical limit of a dual-issue superscalar of IPC=2.

Our new design is loosely coupled, and can couple or decouple its two pipelines dynamically. This paradigm gives the applications the freedom to choose how it should execute based on the need for either high-performance or low-power operation. This feature makes it ideal for embedded applications and cyber-physical domain systems that draw power from batteries or solar panels.

System efficiency equates to reducing operating power. Our new design can achieve an IPC=2 and is suitable to be deployed in an area where power saving is critical, e.g., data center. Besides the power saving, our new design can have the die size reduced by half.

References

[1] M. Sim and Q. Yi, "An Adaptive Multitasking Superscalar Processor," 2019 IEEE 5th International Conference on Computer and Communications (ICCC), Chengdu, China, 2019, pp. 1293-1299.

[2] M. Sim, D. Perera, "A Fast and Secure Pipelined Barrel Processor for Safety-Critical Applications for Real-Time Operating Systems," 2019 IEEE 10th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), New York, NY, 2019.

[3] B. Wheeler. SiFive Raises RISC-V Performance, Series 7 Comprises First Superscalar RISC-V CPUs, The Linley Group, https://sifive.cdn.prismic.io/sifive%2F5ec09861-351b-420c-b6e3-e2b76843044f_linley+report+-+sifive+raises+risc-v+performance.pdf, November 12, 2018

[4] ARM, Cortex-A Series, https://developer.arm.com/ip-products/processors/cortex-a, 2019.

[5] L. Gwennap, Brainiacs, https://www.rose-hulman.edu/class/ee/yoder/ece332/Papers/04%20Brainiacs,SpeedDemons.pdf, Speed Demons, and Farewell, In-Stat MDR, December 27, 1999.

[6] J. Boisseau, L. Carter, K. Gatlin, A. Majumdar, A. Snavely, "NAS Benchmarks on the Tera MTA," Workshop on Multi-Threaded Execution, Architecture, and Compilers, February 1998.

[7] Gaitan, Nicoleta Cristina; Gaitan, Vasile Gheorghita; Ungurean, Ioan; Zagan, Ionel, "Methods to Improve the Performances of the Real-Time Operating Systems for Small Microcontrollers," in Control Systems and Computer Science (CSCS), 2015 20th International Conference on, vol., no., pp.261-266, 27-29 May 2015Apr. 1993.

[8] E. Dodiu and V. G. Gaitan, “Custom designed CPU architecture based on a hardware scheduler and independent pipeline registers – concept and theory of operation,” in IEEE EIT International Conference on Electro-Information Technology, Indianapolis, IN, USA, pp. 1-5, May 2012.

[9] S. Mittal, A Survey of Techniques for Dynamic Branch Prediction, Science and Engineer Research Board (SERB), India, award number ECR/2017/000622, 1st April 2018.

[10] Kevin Skadron, Pritpal S. Ahuja, Margaret Martonosi, and Douglas W. Clark, "Improving prediction for procedure returns with return-address-stack repair mechanisms." MICRO 31 Proceedings of the 31st annual ACM/IEEE international symposium on Microarchitecture Pages 259-271, 1998.

[11] Ravi Hosabettu, Ganesh Gopalakrishnan, and Mandayam Srivas. A Proof of Correctness of a Processor Implementing Tomasulo’s Algorithm Without a Reorder Buffer. In Advances in Hardware Design and Verification: IFIP WG10.5 International Conference on Correct Hardware Design and Verification Methods (CHARME ’99), volume 1703 of LNCS, pages 8-22, Springer- Verlag, 1999.

[12] Ravi Hosabettu, Mandayam Srivas, and Ganesh Gopalakrishnan. Decomposing the proof of correctness of pipelined microprocessors. In Computer-Aided Verification, CAV '98, pages 440-451, 1998.

[13] T. Arons and A. Pnueli. Verifying Tomasulo’s algorithm by refinement. Technical report, Welzmann Institute, 1998.

[14] D. Leibholz and R. Razdan, “The Alpha 21264: A 500 MHz Out-of-Order Execution Microprocessor,” in Proceedings of the 42nd IEEE International Computer Conference, ser. COMPCON ’97. Washington, DC, USA: IEEE Computer Society, 1997, pp. 28, http://dl.acm.org/citation.cfm?id=792770.793692L. Gwennap, Brainiacs, https://www.rose-hulman.edu/class/ee/yoder/ece332/Papers/04%20Brainiacs,SpeedDemons.pdf, Speed Demons, and Farewell, In-Stat MDR, December 27, 1999.

[15] K. Skadron, P. Ahuja, M. Martonosi, D. Clark, "Improving prediction for procedure returns with return-address-stack repair mechanisms," MICRO 31 Proceedings of the 31st annual ACM/IEEE international symposium on Microarchitecture, Pages 259-271, November 1998.

Anomalous Enhancement of Volume and Surface Refractive Index Sensitivity of Fiber Bragg Grating Sensors with Deposition of Gold Nanoparticles

Sankhyabrata Bandyopadhyay, Liyang shao*, Jie Hu, Jiahao Jiang

Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen 518055, Guangdong, China
*Corresponding Author: shaoly@sustech.edu.cn

***Abstract*— in this paper, it has been shown that the refractive index (RI) sensitivity of higher-order cladding modes of a fiber Bragg grating (FBG) can be enhanced significantly with the deposition of gold nanoparticle over the surface. A detailed numerical study is being accomplished in this work. Initially, the sensitivity of higher-order cladding modes is evaluated with a dedicated multi-layer numerical model of FBG. Subsequently, thin nanoparticles of gold are being considered over the surface of FBG. A Four-layer model is employed for the computation of sensitivity of cladding modes of FBG. Almost a '2.5' fold enhancement of sensitivity is found with the utilization of gold nanoparticles. This proposed FBG sensors can be utilized in highly sensitive chemical and biological sensing applications.**

Keywords—Fiber Bragg grating, Gold nanoparticle, Chemical, and Biological sensing

I. INTRODUCTION

Optical fiber-based sensors are being used for selective chemical and biological sensing applications over the past two decades or so. The optical fiber platform of sensing provides label-free, online monitoring, lightweight, and low cost which makes it much more efficient than other chemical and biological sensing applications [1-3]. FBG based sensors are one of the most used and appropriate for highly sensitive chemical and biological sensing applications over other optical fiber sensors for mainly two reasons. Firstly, Wavelength interrogation techniques are being employed in FBG which is always preferable than amplitude-based detection systems and the second reason is that series of FBG can be multiplexed which enables a system for multiparameter sensing [4,5]. Novel biochemical sensors are being developed and reported over the years, like Synchronization of DNA [6], Detection of trace chloroform in water and air [7], antigen/antibody-based immunoassay [8], identification of heavy metal ions in water [9], detection of poisonous gas in the air and many other potential sensing applications [1,2,10,11]. Primarily, after the invention of FBG sensors, it was employed for the detection of strain, temperature, and other physical parameters like velocity, acceleration, force, etc [12, 13]. Very recently, a distributed Bragg reflector-based laser was employed for axial force measurement with a D-shaped fiber [14]. In FBG, forward propagating core mode is being coupled with backward propagation core mode and one reflected wave can be found at the input end whose wavelength of operation is governed by Bragg reflection condition [4,5]. This Bragg reflection condition can be altered with externally applied strain, temperature, force, etc. As the guided core mode didn't interact with the surrounding medium at all, FBG was never used as a RI sensor. The interaction of FBG core mode with the surrounding medium can be enhanced with a significant amount of clad radius. It was demonstrated earlier that with the reduction of clad diameter the volume RI sensitivity of a cladding mode enhanced abruptly [15]. With the reduction, the sensitivity of the cladding mode is being enhanced but the mechanical stability of the sensor is poor and which limits the reuse capacity of the respective sensors [16]. This limitation of the FBG sensor was removed with a slight tilt in the grating plane. The guided core mode of FBG can be coupled with backward propagating core, backward propagating cladding modes, and higher-order radiation modes by a slight tilt in the grating plane. The effective indices of the cladding modes can be altered with a change in surrounding refractive index (SRI) as a result a clear signature can be observed in the spectrum of tilted FBG (TFBG). Distinct chemical and biological sensing applications were accomplished with TFBG [17,18]. The spectrum of TFBG was complex and the fabrication of TFBG with a specific degree of rotation of the grating plane is difficult. Recently, it has been shown that symmetric cladding modes of FBG can be used for volume RI measurement where the sensitivity of the respective cladding mode can be enhanced with mode transition phenomena [19]. The sensitivity of the comparative higher order cladding modes is also being analyzed thoroughly [15, 19], it was demonstrated that the higher-order cladding modes can be employed in highly sensitive biochemical applications. In this work, it has been shown that with the deposition of gold nanoparticles the respective cladding mode can be enhanced significantly. It was demonstrated earlier that the sensitivity of cladding modes of long-period fiber grating (LPFG) was enhanced critically by the deposition of gold nanoparticles over the surface and it was established both experimentally and numerically [20]. It was shown that the surface change of the RI of the gold layer is the basic principle of enhancement of sensitivity for the cladding mode of LPFG [21]. Typically, gold nanoparticles are used as a universal receptor for distinct chemical and biological sensing applications so analysis of the sensitivity of cladding mode with deposition of gold nanoparticles of cladding modes is important [21], in this article we have shown that the sensitivity of cladding mode of FBG in MT effect can be enhanced significantly with deposition of metal nanoparticles. First, the sensitivity of a cladding mode in MT is being computed with a four-layer model. Then a thin gold layer is considered to compute the sensitivity the surface change of RI of the gold layer is taken from previous experimental values

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

observed. Gold nanoparticle coated FBG can be used in novel biochemical applications in the future. The paper will be presented as follows in section II, a concise theory of the structure of the proposed sensor has been discussed. Detail numerical results are being given in section III. Section IV is dealt with discussions and conclusions.

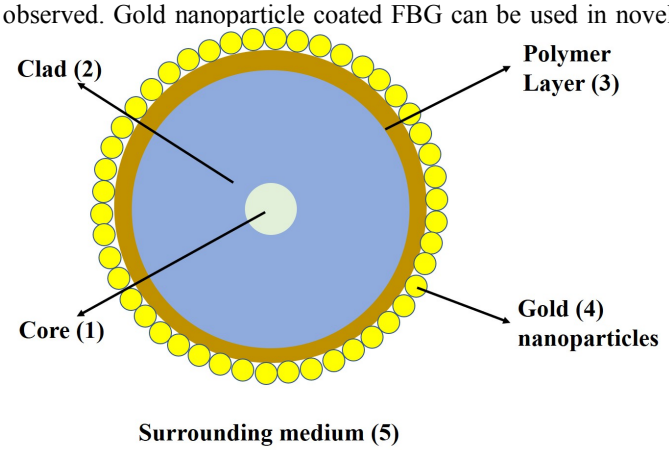


Fig.1. Multi-layer cross sectional diagram of FBG with deposition of gold nanoparticles

II. THEORY

A. Coupled mode equations

In the case of fiber Bragg grating, forward propagating core mode is being coupled with backward propagating core mode and backward propagating cladding modes respectively. They are being governed by the coupled mode equation as shown below [22,23].

$$\frac{dA^{co}}{dz} = ik_{01-01}^{co-co}A^{co} + i\frac{m}{2}k_{01-01}^{co-co}B^{co}\exp(-i2\delta_{01-01}^{co-co}z) + \sum_{v} i\frac{m}{2}k_{0v-01}^{cl-co}B_{ov}^{cl}\exp(-i2\delta_{0v-01}^{cl-co}z) \quad (1)$$

$$\frac{dB^{co}}{dz} = -ik_{01-01}^{co-co}B^{co} - i\frac{m}{2}k_{01-01}^{co-co}A^{co}\exp(-2i\delta_{01-01}^{co-co}z \quad (2)$$

$$\sum_{v}\frac{dB_{0v}^{cl}}{dz} = -i\frac{m}{2}k_{0v-01}^{cl-co}A^{co}\exp(-i2\delta_{0v-01}^{cl-co}) \quad (3)$$

Where A^{co}, B^{co} are the amplitude of forward and backward propagating guided core mode. The term m is the induced modulation. B_{0v}^{cl} are the amplitudes of the backward propagating clad modes of order ν. k_{01-01}^{co-co} is the self-coupling coefficient of the forward and backward propagating core mode. k_{0v-01}^{cl-co} are the cross-coupling coefficients between forward propagating LP_{01} core mode and backward propagating LP_{0v} cladding mode. δ_{01-01}^{co-co} and δ_{0v-01}^{cl-co} are the detuning parameters where

$$\delta_{01-01}^{co-co} = \frac{1}{2}(2\beta_{01}^{co} - \frac{2\pi}{\Lambda}) \quad (4)$$

$$\delta_{0v-01}^{cl-co} = \frac{1}{2}(\beta_{01}^{co} + \beta_{0v}^{cl} - \frac{2\pi}{\Lambda}) \quad (5)$$

They describe the departure from the phase-matching condition, where β_{01}^{co} and β_{0v}^{cl} are the propagation constant of the forward propagating core mode and backward propagating clad mode. Coupling coefficients were computed using the following relation

$$k_{0v,0k} = \int_{\Phi=0}^{2\pi} d\Phi \int_{r=0}^{\infty} \Delta\varepsilon\psi_{0v}(r)\psi_{0k}(r)rdr \quad (6)$$

Where ψ_{0v} and ψ_{0k} are the radial field components of LP_{0v} and LP_{0k} modes. $\Delta\varepsilon$ is the permittivity variation and for this particular waveguide, it is defined as $2\varepsilon n\Delta n$, where Δn is the induced change in the RI of the core during the fabrication of the grating.

The wavefunction of each layer of the multi-layered structure as it is being shown in Fig.1 is evaluated, the wavefunction is normally governed by the equation

$$\psi(r) = A_{v,j}J_v(r\gamma_{v,j}) + B_{v,j}Y_v(r\gamma_{v,j}) \quad \beta_{v,j} < k_0 n \quad (7)$$

$$\psi(r) = A_{v,j}I_v(r\gamma_{v,j}) + B_{v,j}K_v(r\gamma_{v,j}) \quad \beta_{v,j} > k_0 n \quad (8)$$

$J_v(r\gamma_{v,j})$ and $Y_v(r\gamma_{v,j})$ are the ordinary Bessel function of the first and second kind, while $I_v(r\gamma_{v,j})$ and $K_v(r\gamma_{v,j})$ are the modified Bessel function of first and second kind of order ν. The effective indices of the cladding mode are being computed with the transfer matrix method approach [15]. The amplitudes of the core and cladding mode in each of the layers are being evaluated and self and cross-coupling coefficients are computed for the subsequent numerical steps. Finally, a set of the coupled-mode equation is being solved. The coupled model equation can be formed as a matrix and is represented as follows [15,19]:

$$\begin{pmatrix} \frac{dA^{co}}{dz} \\ \frac{dB^{co}}{dz} \\ \frac{dB_{02}^{cl}}{dz} \\ \vdots \\ \frac{dB_{0v}^{cl}}{dz} \end{pmatrix} = \begin{pmatrix} S & \frac{m}{2}S^{co-co} & \frac{m}{2}S_{02-01}^{cl-co} & \cdots & \frac{m}{2}S_{0v-01}^{cl-co} \\ -\frac{m}{2}S^{co-co} & -S & 0 & \cdots & 0 \\ -\frac{m}{2}S_{02-01}^{cl-co} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{m}{2}S_{0v-01}^{cl-co} & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} A^{co} \\ B^{co} \\ \vdots \\ B_{0v}^{cl} \end{pmatrix} \quad (7)$$

where S, S^{co-co}, S_{0v-01}^{cl-co} can be defined as

$$S = ik_{01-01}^{co-co} \quad (8)$$

$$S^{co-co} = ik_{01-01}^{co-co}\exp(-i2\delta_{01-01}^{co-co}z) \quad (9)$$

$$S_{0v-01}^{cl-co} = ik_{0v-01}^{cl-co}\exp(-i2\delta_{0v-01}^{cl-co}z) \quad (10)$$

The variation of effective indices with the deposition of overlay material is being analyzed with a five-layer numerical model. The detailed description of numerical work is being presented in the following sections.

III. NUMERICAL RESULTS

The parameters of a standard single-mode fiber SMF28-e is being considered. The refractive indices of the core and the clad are n_1=1.4494 and n_2=1.444 (@ 1550 nm). Diameters of the fiber core and the clad were considered as 8.25 µm and 125 µm respectively. Recently it was shown that the comparative higher order cladding modes possess exceptional MT characteristics along with the high sensitivity. Initially, the spectrum of the FBG is computed with considering the first 45 cladding modes with a three-layer numerical model and which is being shown in Fig.2. The spectrum is being computed with a three-layer model, namely a core-clad-surrounding medium. The period of the grating was ~ 0.576 µm. The spectrum is

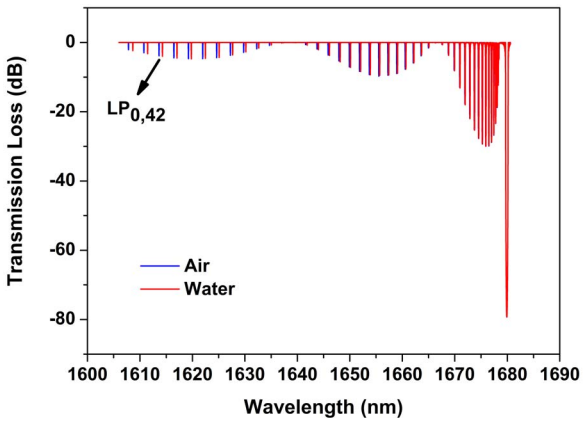


Fig.2. Numerically computed spectrum of FBG considering the surrounding medium as air and water

being computed with consideration of the surrounding medium like air (RI ~ 1) and water (RI ~ 1.333). The $LP_{0,42}$ cladding mode is also identified in Fig.2. Then MT characteristics of $LP_{0,41}$ and $LP_{0,42}$ cladding mode is being shown in Fig.3a and as a comparison, the MT effect the lower order cladding modes (LP_{06} and LP_{07}) is being depicted in Fig.3b. The RI of the overlay has been considered to be ~ 1.7, which is a standard RI of Silica-Titania sol-gel materials. It is to be noted that the MT behavior of LP_{06} and LP_{07} cladding mode is completely different than $LP_{0,41}$ and $LP_{0,42}$ cladding modes. The optimum overlay thickness for LP_{06} and LP_{07} cladding mode is being shown in Fig.3a. The slope of the variation of effective indices with overlay thickness is being maximum at this point. A thickness of ~ 150 nm of an overlay is required to reach the most sensitive point in the case of LP_{07} cladding mode but few tens of nm of overlay thickness are adequate to reach the most sensitive point for $LP_{0,42}$ cladding mode. The scenario of the MT effect is different in the case of higher-order cladding modes, the slope of the graph is quite steep from the beginning of the deposition of the overlay. A detailed discussion of the sensitivities of higher-order cladding modes of FBGs is being illustrated elsewhere [15]. The analysis of add-layer sensitivity was also being accomplished in that work as it was established that surface-based measurements cannot be quantified with the concept of volume RI sensitivity. The add-layer sensitivity concept was employed successfully for LPFG sensors [24]. In this work the volume and surface RI sensitivity of cladding modes of FBGs are being computed with deposition of gold nanoparticles, for the modeling, we have considered a thin layer of gold layer of 10 nm and it was shown that with a high surface coverage of deposited gold nanoparticles average thickness is closely approached to the diameter of gold nanoparticles [19]. The deposited gold nanoparticles have some intermediate spaces and clusters of nanoparticles can be observed, so the RI of the gold layer is a composite of characteristics RI of gold and the RI of the medium. As a consequence, if the RI of the external medium is being changed the RI of the Gold layer is also changed and it has been established with theory and experiments. In this article, the change in surface and volume RI sensitivity with the deposition of gold nanoparticles is being analyzed. First, in the case of LP_{07} cladding mode, 135 nm of overlay and then 10 nm of the gold layer was considered after a polymer layer of ~ 20 nm [20] to reach the maximum sensitive region, it is

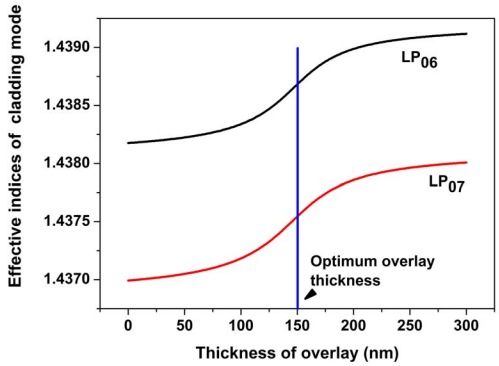


Fig.3a. Mode transition phenomena of LP_{06} and LP_{07} cladding mode with deposition of overlay

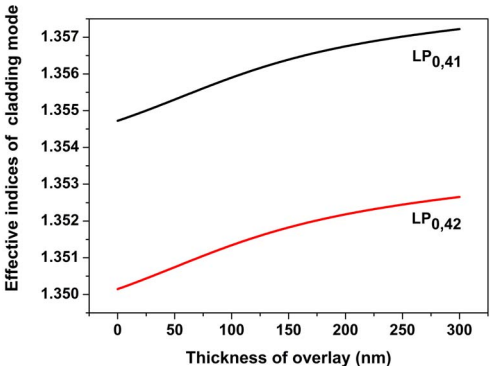


Fig.3b. Mode transition phenomena of $LP_{0,41}$ and $LP_{0,42}$ cladding mode with deposition of overlay

observed that without the gold layer the volume RI sensitivity of LP_{07} cladding mode is ~ 11,200 pm/RIU at the highest sensitive point. The RI of the external medium is very near to the water, the sensitivity is being computed with an RI zone of 1.3333 to 1.3433. The volume RI sensitivity with the deposition of the gold layer is found to be ~ 25,300 pm/RIU within the same RI zone of the external medium. This simulation was accomplished with the five-layer architecture of FBGs namely core-clad-overlay/polymer layer-gold layer-surrounding. With the same method, the sensitivity of $LP_{0,42}$ cladding mode was studied and 20 nm of polymer layer (PDDA with RI 1.54) and 10 nm gold layer is being considered

for the computation purpose. The volume RI sensitivity with and without gold nanoparticle is 35760 pm/RIU and 74380 pm/RIU respectively. It is being observed that with the deposition of the gold layer the sensitivity is being enhanced significantly. The computed shift of the spectrum of $LP_{0,42}$ cladding mode is being given in Fig.4. Further, the add-layer sensitivity is being computed with the extension of the numerical model to eight layers namely "core-clad-overlay-gold layer-receptor-target-surrounding medium". In this particular case, the thickness of the overlay is considered to be ~ 100 nm, the gold-layer is of 10 nm, the receptor is 10 nm and the target is 20 nm, the RI of the biolayer is being chosen as ~ 1.56 [23] for the computation of the add-layer sensitivity of LP_{07} cladding mode. The thickness of the overlay of $LP_{0,42}$ cladding mode is selected as ~ 20 nm. The computed add layer

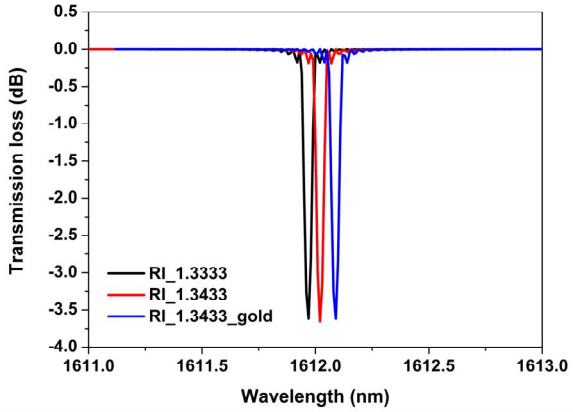


Fig.3b. Computed shift of $LP_{0,42}$ resonance cladding mode with change in SRI from 1.3333 to 1.3433, RI_1.3333_gold means shift of respective cladding mode with consideration of gold nanolayers in the surface.

sensitivity of LP_{07} cladding mode with and without gold layer is founded to be ~ 17 pm/nm attachment of biolayer and 28 pm/ nm attachment of the biolayer respectively. The add layer sensitivity has been computed with keeping the external medium fixed at PBS medium (i.e RI ~ 1.334) respectively. Whereas, the add-layer sensitivity of $LP_{0,42}$ cladding mode is ~ 46 pm/ nm attachment of biolayer and 68 pm/ nm attachment of biolayer with deposition of gold nanoparticles.

IV. Discussion and Conclusion

In this paper, detailed numerical work is being accomplished to study the enhancement of volume and surface RI sensitivity of lower and higher-order cladding modes of FBG. It is to be mentioned that, MT effect is also being amalgamated with the effect of gold nanoparticles. For the evaluation of sensitivity multi-layer models of FBG is being considered and which has been explained in this article. The resolution of RI detection is ~ $1x10^{-6}$ with standard FBG interrogator with 1 pm resolution which is quite significant. Interestingly with the deposition of gold nanoparticles at-least, '2.5' fold enhancement of the sensitivity is being observed. The concept of effective medium RI is being considered for the evaluation of sensitivity. As gold nanoparticles is being used extensively for the universal receptors for distinct bio and chemical target analyte this concept of enhancement of sensitivity can open up a new avenue for designing highly sensitive optical fiber sensors for selective biochemical applications.

Acknowledgment

This work is supported by the Post-doctoral research grant and a start-up grant from the Southern University of Science and Technology, Shenzhen, China.

References

1. S. Pissadakis "Lab-in-a-fiber sensors: A review", Microelectronics engineering, vol. 217, 111105, 2020.
2. F. Chiavaioli, F. Baldini, S. Tombelli, C. Trono and A. Giannetti "Biosensing with optical fiber gratings" Nanophotonics 2017; 6(4): 663–679.
3. X Wang and O. S. Wolfbeis, "Fiber Optic Chemical sensors and biosensors (2013-2015), "Analytical Chemistry, vol. 88, pp.203-227, 2016.
4. G. Allwood, G. Wild and S. Hinckley "Fiber Bragg Grating Sensors for Mainstream Industrial Processes", MDPI Electronics, vol. 6, pp.1-19, 2017.
5. C. Broadway, R. Min, A. Gomes, C. Marques, and C. Caucheteur, "Toward Commercial Polymer Fiber Bragg Grating Sensors: Review and Applications" IEEE Journal of Lightwave Technology, vol.37, no.11, pp. 2605-2615, 2019.
6. X. Chen, L. Zhang, K. Zhou, E. Davies, K. Sugden, I. Bennion, M. Hughes, and A. Hine, " Real-time detection of DNA interactions with long period fiber-grating-based biosensor, " Optics Letters, vol.32, pp. 2541-2543, 2007.
7. M. Giordano, M. Russo, A. Cusano and A. Cutolo G. Mensitieri and L. Nicolais, "Optical sensor mbased on ultrathin films of -form syndiotactic polystyrene for fast and high resolution detection of chloroform," Applied Physics Letters, vol.85, pp.5349- 5351, 2004.
8. S. Maguis, G. Laffont, P. Ferdinand, B. Carbonnier, K. Kham, T. Mekhalif and M. Millot, "Bio-functionalized tilted Fiber Bragg Gratings for label-free immunosensing" Optics Express, vol. 16, pp. 19049-19054. 2008.
9. J. Yanga, L. Chen, Y. Zheng, X. Dong, R. Raghunandhan, P. Soe, C. Chan, "Heavy metal ions probe with relative measurement of fiber Bragg grating" Sensor Actuator B, vol. 230, pp. 353-358, 2016.
10. Sridevi. S, K. Vasu, N. Bhat, S. Asokan, A. K. Sood, "Ultra sensitiveNO2 gas detection using the reduced graphene oxide coated etched fiber Bragg gratings," Sensor and Actuator B, vol. 223, pp. 481-487, 2016.
11. C. Zhang, S. Xu, J. Zhao, H. Li, H. Bai, and C. Miao, " Differential intensity modulation refractometer based on SNS structure cascaded two FBGs," IEEE Photonics Journal, vol. 9 pp. 7103008, 2017.
12. B. Guan, H. Y. Tam, X. M. Tao, and X. Dong "Simultaneous Strain and Temperature Measurement

Using a Superstructure Fiber Bragg Grating" IEEE Photonics Technology Letters, vol. 12, pp. 675-677,2000.

13. S. Bandyopadhyay, L. Shao, W. Chao, Z. Yan, F. Hong, G. Wang, J. Jiang, P. Shum, X. Hong, W. Wang" Highly efficient free-space fiber coupler with 45° tilted fiber grating to access remotely placed optical fiber sensors", Optics Express, vol. 28, pp. 16569-16578. 2020.
14. X. Yang, S. Bandyopadhyay, L.-Y. Shao, D. Xiao, G. Gu, and Z. Song, "Side-Polished DBR Fiber Laser with Enhanced Sensitivity for Axial Force and Refractive Index Measurement", IEEE Photonics Journal, vol. 11, pp. 1-10, 2019.
15. S. Bandyopadhyay, L. shao, M. Smietana, C. Wang, J. Hu, G. Wang, W. He, G. Gu, Y. Yang "Employing higher order cladding modes of fiber Bragg grating for analysis of refractive index change in volume and at the surface" IEEE Photonics Journal, vol.12 pp. 7100313, 2020.
16. W. Liang, Y. Huang, Y. Xu, R. K. Lee, and A. Yariv, "Highly sensitive fiber Bragg grating refractive index sensors," Applied Physics Letters, vol. 86, no. 15, pp. 151122-1–151122-3), 2005.
17. G. Laffont and P. Ferdinand, "Tilted short-period fibre-Bragg-grating-induced coupling to cladding modes for accurate refractometry," Measurement of Science and Technology. vol. 12, pp-765-770, 2001.
18. J Albert, L Shao, C Caucheteur, "Tilted Fiber Bragg grating Sensors," Laser and Photonics Review., vol.7, pp. 83-108, 2013.
19. S. Bandyopadhyay, T. Dey, N. Basumallick, P. Biswas, K. Dasgupta, and S. Bandyopadhyay "High Sensitive Refractometric Sensor Using Symmetric Cladding Modes of an FBG Operating at Mode Transition" IEEE Journal of Lightwave Technology, vol.34, no.14, pp. 3348-3353, 2016.
20. S. Bandyopadhyay, N. Basumallick, S. Bysakh, T. Dey, P. Biswas, S. Bandyopadhyay, "Design of turn around point long period fiber grating sensor with Au-nanoparticle self monolayer" Journal of Opt. Laser. Tech. 102, (2018), 254.
21. K. Saha, S. Agasti, C. Kim, X. Li, V. Rotello, "Gold nanoparticles in chemical and biological sensing", Chemical Reviews. Vol. 112, pp. 2739–2779, 2012.
22. T. Erdogan, and J. Sipe, "Cladding-mode resonances in short- and long period fiber grating filters," Journal of Optical Society of America A, vol. 14 , pp. 1760-1773, 1997.
23. E. Anemogiannis, E. N. Glytsis, and T. K. Gaylord," Transmission characteristics of long-period fiber gratings having arbitrar azimuthal/radial refractive index variations," IEEE Journal of Lightwave and Technology, vol. 21, pp. 218-227, 2003.
24. S. Bandyopadhyay, I. Villar, P. Biswas, N. Basumallick and T. Dey, S. Bandyopadhyay," Long Period Fiber Grating for Biosensing: An Improved Design Methodology to Enhance Add-Layer Sensitivity " IEEE Journal of Lightwave Technology, vol. 36, pp. 1178-1184, 2018.

Deployment Optimization of Dynamic Wireless Electric Vehicle Charging Systems: A Review

Eiman A. ElGhanam, Mohamed S. Hassan, and Ahmed H. Osman
Department of Electrical Engineering
American University of Sharjah
P.O. Box 26666, Sharjah, UAE
{eelghanam, mshassan, aosmanahmed}@aus.edu.

***Abstract*—To alleviate the range anxiety fear of electric vehicle (EV) drivers, dynamic wireless charging (DWC) systems are being developed to supply energy to the EV during its motion, thereby compensating for EV energy consumption and extending the vehicle's driving range. Nevertheless, several challenges are involved in the commercial deployment of on-road DWC systems, particularly with respect to the associated deployment costs and EV energy demand. Accordingly, extensive research is conducted to optimize the deployment of DWC systems within a city infrastructure while addressing the different challenges to encourage mass adoption of EVs and improve their reliability. In this work, a review of current state-of-the-art research into optimizing the deployment of dynamic EV wireless charging facilities is presented, to provide a set of guidelines on the formulation and optimal deployment of DWC systems. In addition, on-the-move V2V energy exchange is also addressed as an additional dynamic charging solution to complement on-road charging systems.**

***Index Terms*—Electric vehicle, dynamic wireless power transfer, lane deployment, location optimization, charging cells.**

I. Introduction

In order to reduce the range anxiety of electric vehicle (EV) users and promote mass adoption of this environment-friendly means of transportation, dynamic wireless charging (DWC) systems are gaining an increasing global interest. This is because DWC systems enable EVs to compensate for their consumed energy during motion without having to stop for frequent recharging and/or battery swapping. This accordingly increases the maximum non-stop trip mileage of the EV, extends its driving range and prevents excessive depletion of the EV batteries thereby prolonging the battery lifetime [1]–[3]. The most common structure of a DWC system utilizes resonant inductive power transfer (RIPT) to transfer electrical energy from a set of primary coils buried under the ground to a secondary coil fitted at the bottom of the EV, separated by an air-gap equivalent to the vehicle-to-ground clearance distance. This is also referred to as Grid-to-Vehicle (G2V) power transfer. In order to ensure optimum energy management and integration of EVs into the city infrastructure, the RIPT circuit level design needs to be complemented with a city-wide deployment optimization framework, to maximize the benefits of this dynamic charging technology. In particular, optimal deployment of EV DWC systems aims to achieve one or more of the following objectives:

- Maximizing the energy received by the EVs in the network,
- Minimizing infrastructure deployment costs,
- Enhancing EV battery performance and lifetime,
- Optimizing charging power levels,
- Minimizing the number of infeasible trips due to EV battery capacity limitations.

Accordingly, several deployment optimization studies have been reported in the literature, of which the On-Line Electric Vehicle (OLEV) project demonstrates one of the earliest commercial deployments [4]–[7]. Other deployment optimization models are also reported in the literature, including driving-cycle based allocation models [8]–[10], traffic flow-based analysis [11]–[13] and deployment at signalized intersections [14], [15]. Vehicle-to-vehicle (V2V) dynamic charging solutions using mobile energy disseminators (MEDs) are also proposed in [16], [17], in which RIPT is employed to dynamically transfer energy between EVs to enable the energy-demanding vehicles to complete their trips without the need to stop for prolonged charging durations. In this work, the authors present a review of current state-of-the-art research into the deployment of dynamic G2V and V2V EV wireless charging solutions, aiming to summarize the different deployment optimization strategies and present a set of guidelines to enable researchers and city planners to formulate and determine the most optimum facilities deployment plan(s).

The rest of this paper is organized as follows: Section II describes the structure of the G2V DWC system addressed in this work in order to accurately define the deployment optimization framework. Section III then presents a detailed review of state-of-the-art G2V DWC system deployment optimization models. This is then followed by a description of V2V dynamic energy exchange solutions in Section IV before the paper is finally concluded in Section V.

II. G2V DWC System Model

The RIPT-based dynamic wireless EV charging system consists of two physically-separated coils that exchange power using alternating magnetic fields, with no direct electrical

This work is jointly supported by the American University of Sharjah through SCRI Grant No. SCRI 18-CEN-10 and by Sharjah Research Academy.

978-1-7281-9615-2/20/$31.00

connection between them. The primary coil laid on the ground is connected to the mains grid through rectification, power factor correction and high frequency inversion circuitry. These work together to generate high-frequency alternating magnetic fields that are coupled with the secondary coil fitted at the bottom of the EV. This enables wireless power transfer from the grid to the vehicle; i.e. Grid-to-Vehicle (G2V) power transfer, as shown in Figure 1.

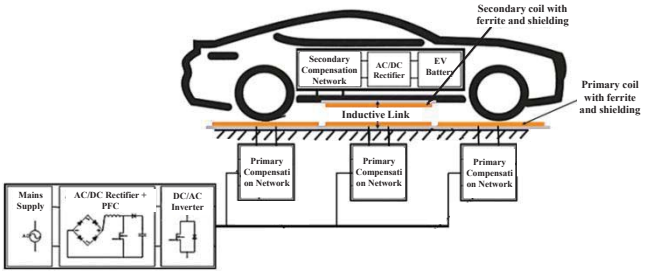


Fig. 1: Dynamic EV charging system using G2V power transfer.

The power received by the EV is then rectified, regulated and supplied to the EV battery to compensate for the energy consumed during the vehicle's motion while contributing to the EV's energy reserve. The relationship between the energy consumed by the EV, the received energy from DWC and the State-of-Charge (SoC) of the EV battery can be simplified as:

$$SoC_f = SoC_i + \left(\frac{E_{DWC} - E_{con}}{E_{max}}\right) \times 100, \qquad (1)$$

where SoC_f is the final SoC at the end of the trip, SoC_i is the initial SoC, E_{DWC} is the energy received from the dynamic charging system, E_{con} is the amount of energy consumed by the EV over the same trip, and E_{max} is the maximum EV battery capacity. The received energy, E_{DWC}, is related to the grid power supply, P_{In}, through the efficiency of the DWCS system, η_{DWC} using:

$$E_{DWC} = \int_{\mathbb{C}} \eta_{DWC} P_{In}(t) dt, \qquad (2)$$

where $\mathbb{C}$ is the time over which the EV is receiving energy from the primary charging coils. The efficiency of the DWC system , and hence the energy received by the EV are both dependent on the design of the different system components shown in Figure 1, including the design of the primary and secondary coils and their respective compensation networks, and are impacted by the coupling performance of the inductive link and the vehicle's alignment over the primary coils [18]. In addition, accurate estimation of the SoC of the EV battery needs to acknowledge battery temperature, aging, and other factors that affect the battery energy levels. Complete details of the RIPT DWC system design and EV battery SoC estimation methods can be extensively found in [2], [19]–[21] but are beyond the scope of this paper. In this work, a high-efficiency DWC system is assumed to be readily available and the problem at hand is to ensure optimum implementation and deployment of this system within a city infrastructure.

III. State-of-the-Art G2V DWC System Deployment Optimization Models

The optimal DWC system deployment problem can be classified into three main categories based on the corresponding solution space:

- **Macro allocation model**, for city-wide deployments in which standardized driving cycles, such as Urban Dynamometer Driving Schedule (UDDS) and the HighWay Fuel Economy Test (HWFET) [22], and/or metropolitan road network plans are utilized to determine the optimum location of charging coils.
- **Micro allocation model**, in which optimum routes between origin-destination (O-D) pairs are selected for the deployment of charging coils based on route feasibility and traffic conditions.
- **Deployment on signalized intersections**, by utilizing the relatively lower EV speeds at intersections and red-light stopping times to maximize the energy received by the EV. This is also referred to as *quasi-dynamic charging*.

Accordingly, the references highlighted in the introduction of this paper are classified based on these categories and their optimization solutions are discussed in the remainder of this section.

A. Macro Allocation Model

In the macro allocation deployment optimization problem, information on EV speeds and acceleration profiles is obtained from standardized driving cycles over predefined durations, and is used to recommend charging coil deployments at areas of slowest speeds, higher power requirements, etc. In fact, standardized driving cycles, also known as driving schedules, are typically used for vehicle emissions and fuel economy testing due to the nature of the information they carry, yet can be adopted for deployment optimization of DWC systems. This is demonstrated by the authors in [8], [23], in which the UDDS (low power demand, urban driving cycle), as well as other standardized cycles, have been used to determine the optimum locations of the charging coils. A speed-time plot of the UDDS is shown in Figure 2.

As observed in Figure 2, the UDDS cycle has a low average speed of around 31.5 km/hr as well as several zero-speed slots in which EVs are expected to be idle, possibly at intersections or during congestion periods. Mapping these zero-speed durations to a city road network suggests the deployment of charging coils at these locations to maximize the received energy by the EV. This is highlighted in [8], in which the author concludes that multiple 10 m-long primary coils are required to cover up to 20% of a UDDS cycle with $P_{In} = 40$ kW in order to extend the driving range of a typical EV by around 87% while maintaining the battery SoC above 20%, assuming that the initial SoC is 100%. On the other hand, the authors in [23] choose to allocate a single 500 m long lane of primary coils to the lowest speed section of

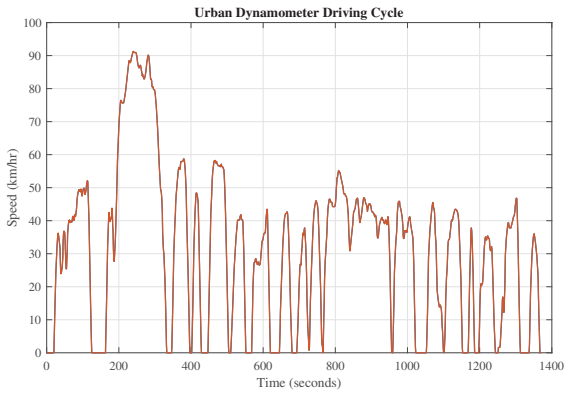


Fig. 2: Speed-time plot of UDDS cycle.

the UDDS cycle to achieve an 85.7% extension in driving range. In addition, the authors in [9] further extended the macro allocation model by including the temporal aspect of the deployment process together with the spatial allocation. This is addressed by incorporating the road lifetime assessment model into the optimization framework, in order to determine the most optimum year at which a particular road can be modified to include charging cells, based on current and expected EV sales and required cost of infrastructure for DWPT system deployment.

Another city wide deployment plan is addressed in [10] in which the authors used mobility records of all public transit vehicles within the city to define the solution space, then design for the optimum deployment that minimizes the total deployment cost by minimizing charging lane length and reducing the number of deployed lanes, while ensuring minimum residual energy is available in the EV battery at each node in the network. The objective function is as follows:

$$\textbf{Minimize Cost} = \sum_{i \in N} \omega_0 x_i L_i, \tag{3}$$

subject to:

$$SoC_i \geq \alpha, \tag{4}$$

where N is the set of candidate locations, ω_0 is the unit cost, L_i is the length of the charging segment at location i, x_i is a binary variable to determine whether or not the location has a charging segment, and α is the residual energy threshold.

The authors started with a categorization and clustering process, in which they identified and grouped potential charging cell locations based on a set of attributes, namely vehicle passing speed and visit frequency for each location. A ranking is then allocated to each cluster and candidate locations are identified. Optimum charging cell locations are then selected from the candidate locations such that they provide minimum total deployment cost while ensuring a certain level of expected residual energy is maintained in the EV battery at each location. This energy level is the minimum energy required to reach the next charging location. The clustering-based approach utilized by the authors in [10] helps eliminate unsuitable locations and hence, reduces the solution space of the optimization problem and further simplifies the analysis.

A subset of the city-wide macro allocation model is the deployment optimization for a set of predefined routes with known speed profiles, which is particularly applicable to public transportation electric buses. This is demonstrated by the authors in [4] and [5] to optimize the deployment of wireless charging lanes along the route of the OLEV bus, and is further extended by the authors in [6] and [7] to address a multi-route environment. The optimization objective function is to minimize the total cost of the OLEV project deployment, as shown in the following expression [4]:

$$\textbf{Minimize Cost} = kc_b E_{max} \frac{T_t}{T_b} + c_f N + c_v \sum_{i=1}^{N} (x_i^e - x_e^s), \tag{5}$$

where E_{max} is the maximum battery capacity, T_t is the lifetime of the charging track, T_b is the battery lifetime, c_b is the fixed battery cost, c_f is the fixed cost per charging section, c_v is the variable cost depending on the length of the section, N is the number of charging sections, and x_i^s and x_i^e are the starting and ending points of each charging section, respectively.

The cost function in (5) consists of battery cost based on the battery lifetime, fixed infrastructure costs related to the cost of the inverters and other power electronics circuitry, and variable infrastructure costs related to the length of the charging sections. Accordingly, the optimum location of the wireless charging sections is determined along with the optimum bus battery size, using the meta-heuristic Particle Swarm Optimization (PSO) algorithm. A similar optimization scenario is also tackled by the authors in [24] using the heuristic Genetic Algorithm (GA) instead of PSO. In their work, the authors utilized a detailed power consumption model of shared automated electric shuttles (SAES), and expanded the battery cost coefficient c_b into a function of the charging rate of the battery, i.e. C-rate, and the operating SoC window during typical shuttle bus operation on the assigned route. Accordingly, their optimization process revealed that few high-power 100 kW wireless charging segments can be placed along the pre-determined route to provide the required charge-sustained operation with zero charging downtime and reduced battery size by around 36%. The joint optimization of charging coil deployment and bus battery sizing is also addressed in [25] for an eight-route bus network, while including the uncertainty of bus energy consumption and battery depletion rate in a Robust Optimization (RO) model to improve the accuracy of the optimization outcome.

EV battery downsizing, while important for the public transportation network, is difficult to be included in a city-wide optimization model that involves privately-owned EVs with different EV battery capacities and different energy consumption patterns. Accordingly, another city-wide optimization model based on cost minimization is addressed in [26] without attempting to downsize EV batteries. In this work, the author aimed to minimize the capital costs of dynamic

charging infrastructure implementation to enable an EV to travel between key destinations in California State, subject to constraints on battery capacity, reserve mileage requirements, and vehicle charging levels. Different combinations of wireless charging power (up to 120 kW) and vehicle range (up to 300 miles) are investigated and the author concludes that to achieve a driving range of 200 miles, a DWC system with 100 kW charging power costs $1 billion less than a 40 kW DWC system. The authors in [27] also define a minimum infrastructure cost minimization objective function and acknowledge the DWC system power ratings yet do not incorporate any battery-related details in the presented analysis.

B. Micro Allocation Model

In contrast to the citywide macro allocation models, other researchers chose to address DWC system deployments in a microscopic scale, by defining the solution space into a set of routes between an origin (O) and a destination (D) and acknowledging traffic conditions along these routes. An O-D road network can be modeled as a directed graph, $\mathfrak{G}$, consisting of $\mathfrak{M}$ nodes and $\mathfrak{L}$ links joining a set of origins $\mathfrak{O}$ to a set of destinations $\mathfrak{D}$. A sample directed graph with a single O-D pair and $M = 3 \times 3$ nodes is shown in Figure 3. The location and direction of the links between each pair of

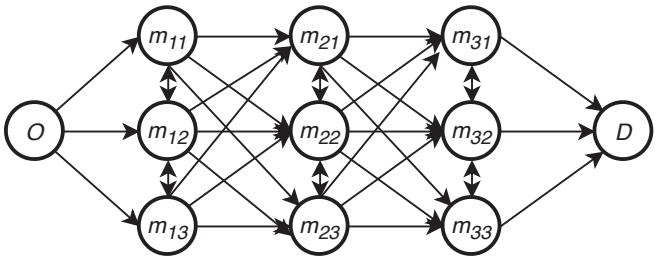


Fig. 3: Sample directed road network graph with a single O-D pair and nine interconnected nodes.

adjacent nodes shall reflect the actual paths in the network. Accordingly, the deployment optimization problem is solved by route optimization, such that the charging sections are placed on the most optimum links to minimize the number of infeasible routes between an O-D pair. A feasible route is defined as a complete path from the origin to the destination that can be completed by an EV while maintaining its SoC above a predefined threshold. The authors in [11] aimed to maximize the number of feasible routes between an O-D pairs subject to a given budget, and define two optimization problems for city planners:

1) For a given minimum EV energy threshold, determine the minimum budget of charging lane deployment, together with the corresponding locations, needed such that the number of infeasible routes is zero. This is equivalent to:

$$\textbf{Minimize Cost} = \sum_{l=1}^{L} c_l x_l, \tag{6}$$

where c_l is the cost of installation of a charging unit along a link, x_l is a binary variable to indicate whether or not the charging unit is installed and L is the total number of links between the O-D pair assuming all routes are feasible.

2) For a given minimum energy and budget constraints, determine the optimal installation locations to minimize the number of infeasible routes and hence reduce the range anxiety problem of drivers within the network. This is equivalent to:

$$\textbf{Maximize Feasible Routes} = \sum_{r=1}^{R} \delta_r w_r y_{lr}, \tag{7}$$

where r is the index of the route (consisting of multiple links L), δ_r is the normalized travel demand along the route, w_r is a binary variable to indicate whether or not the EV completes the route with its SoC above the defined threshold and y_{lr} is a binary variable to indicate whether or not link l belongs to route r. The authors use δ_r to omit road segments that have low traffic demand from the set of feasible routes, thereby reducing the solution space of the optimization model.

For both problems, the authors in [11] performed a computational evaluation of the objective functions for a given directed graph with 26 nodes using a simplified SoC estimation similar to that in (1). Results revealed significant improvements, i.e. reduced deployment costs and increased feasible routes, in contrast to other heuristics that forcefully install charging units in mostly visited links and/or in centralized links.

On the other hand, the authors in [12] chose to allocate charging sections such that the trip time is minimized between an O-D pair while ensuring that the EV completes the trip without fully depleting its battery. The authors in [13], however, located their charging sections based on an objective function that maximizes the captured flow, i.e. maximizes the number of paths that an EV can take between an O-D pair without running out of energy. The interaction between traffic flow patterns and the location of the charging facilities is incorporated by applying the stochastic user equilibrium principle to describe the routing choice behavior of EV drivers based on the availability of charging cells. This is also acknowledged in [12] in which network equilibrium between traffic flow and demand is set as a constraint to ensure that all feasible paths are utilized and considered for charging lane deployments. Accordingly, both [12] and [13] aimed at placing the charging sections at locations of maximum traffic flow to ensure maximum utilization of the charging infrastructure.

C. DWC System Deployment at Signalized Intersections

As the macro allocation model recommends regions of slowest speed for charging lane deployment, traffic intersections are strong candidates for the optimal deployment problem. Nevertheless, the microscopic model emphasizes the importance of route feasibility, which is assessed based on the traffic flow and other route conditions between O-D

pairs. Accordingly, an integration of both allocation models is evident in quasi-dynamic charging system deployments at signalized intersections, in order to benefit from the relatively lower vehicle speeds and pre-known traffic conditions to maximize the energy received by the EVs while controlling traffic signal timings.

Optimal deployment at signalized intersections is addressed by the authors in [14] where they proposed an optimal combination of traffic signal timings and wireless charging infrastructure locations based on traffic conditions to maximize the charging energy while reducing traffic delay by controlling signal timings. In order to model the road network and incorporate traffic conditions, the modified cell transmission model (CTM) is used where a lane is discretized into a number of cells of equal width and the cells are classified based on their location on the signalized intersection. This is shown in Figure 4. Accordingly, each cell is assigned a

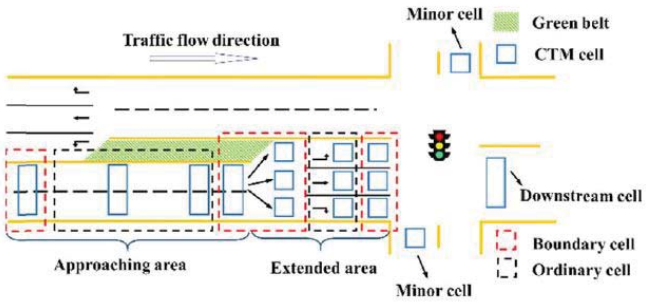


Fig. 4: Modified cell transmission model (CTM) adopted in [14].

binary variable to determine whether or not charging coils are required, and traffic flow in and out of each cell is used to determine the optimum signal timings. The optimization model is hence a mixed integer, non-linear programming (MINLP) model with continuous and discrete constraints. The authors started by choosing a set of feasible signal timings then using traffic flow and traffic density information with signal timings to solve a binary programming model to determine charging cell locations. Accordingly, the authors filtered these locations to create a Pareto set of possible solutions and eliminated solutions with larger delay and smaller energy, ultimately reaching a single optimal solution. By utilizing traffic flow and density information, the authors acknowledged the distribution of vehicles along the road and aimed to meet their energy demand by adjusting traffic signal timings with optimal charging cell locations.

On the other hand, without utilizing the CTM, the authors in [15] also addressed quasi-dynamic charging at signalized intersections, by defining their optimization problem to maximize the total energy transferred to all EVs by optimizing the green signal times and the length of the charging lanes placed at the intersections. However, instead of incorporating actual traffic flow data in a hybrid genetic and PSO optimization algorithm similar to the authors in [14], the authors in [15] utilized traffic simulations and queue counting techniques on the Simulation of Urban MObility (SUMO) simulator [28] to obtain information on received energy for different lane lengths and green signal timings and used it to define the solution space for the optimization objective function.

In contrast to controlling the signal timings, the authors in [29] utilized existing red traffic light durations to determine the Waiting Queue Length (WQL) and use that to represent the charging demand of EVs at the intersection. On the other hand, the charging supply is represented by the Wireless Charging lane Length (WCL) and the optimization problem aimed to determine the optimal length of charging lanes that need to be deployed to minimize the charging lane installation costs while minimizing the gap between charging energy supply and demand. Accordingly, the authors used real traffic datasets to define the charging demand in an urban area and build the corresponding charging lane deployment plan to meet this demand while ensuring minimum construction costs.

IV. V2V DWC System Deployment Model

The on-the-move V2V energy exchange is another approach proposed in the literature to reduce range anxiety of EV drivers by offering a dynamic charging solution. This utilizes city busses, with high-capacity batteries, to wirelessly supply power to vehicles in demand for charging energy, thereby operating as Mobile Energy Disseminators (MEDs). In this model, the primary coil is fitted in front of the MED bus and the secondary coil is at the back of the EV. RIPT is employed to enable the energy exchange between the two vehicles. This is shown in Figure 5.

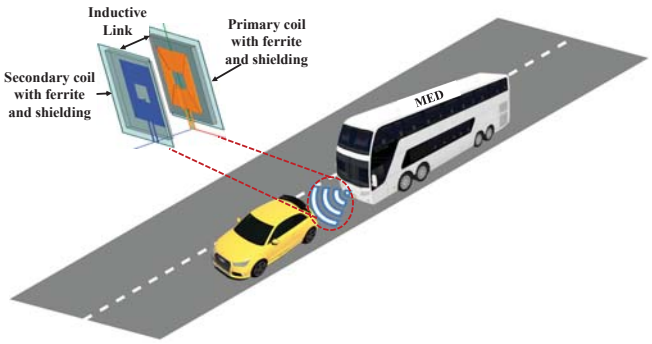


Fig. 5: Dynamic EV charging system using V2V power transfer.

In [16], the authors utilized mobile energy disseminators (MEDs) to optimize the route taken by EVs to reach from source to destination. Accordingly, the optimization problem is a restricted shortest path problem in which the optimum path from Node A to Node B is selected based on the presence of MEDs on a road segment where the additional distance required to meet MEDs and the predicted total travel time are minimized. The weight function used for path selection then consists of three main components: travel distance, travel time and EV energy consumption along the path. The optimization

objective is to minimize the weight, W assigned to each road segment j, along the route of the EV i, as follows:

$$\textbf{Minimize} \sum_{ij} W_{ij}, \tag{8}$$

in which

$$W_{ij} = F(E_{ij}, T_{ij}, D_{ij}), \tag{9}$$

where E_{ij} is the energy consumption of vehicle i along road segment j, T_{ij} is the corresponding travel time and D_{ij} is the corresponding travel distance.

On the other hand, the authors in [17] also utilized MEDs for route optimization by minimizing the total travel time for every EV while utilizing both MEDs and static charging stations. Yet, they did not consider minimizing the travel distance or energy consumption as an objective of the optimization problem. Instead, their work focused on minimizing the travel time by acknowledging the additional time required to use the path that has an MED as well as the charging time and the waiting time on static charging stations. The EV energy level was used as a constraint that determines whether the path is feasible or not.

The concept of Mobile Energy Disseminators requires effective V2V energy exchange and accurate coordination between the MED and the EV to maintain front-to-rear distance and avoid accidents. Nevertheless, the utilization of MEDs helps to minimize the number of infeasible trips by addressing urgent on-the-move EV energy demand without the need for excessive investments in infrastructure transformation to accommodate the charging lanes. MEDs are hence expected to offer faster time-to-market (TTM) than deployments of DWC systems, although challenges in authentication, billing and effective energy distribution may be faced by the MED operators.

V. Conclusions

In this work, a review of the deployment optimization frameworks of different dynamic wireless EV charging solutions are presented. On one hand, different on-road charging unit deployment optimization models are described and categorized based on the solution space of the optimization problem and the associated variables. State-of-the-Art research findings into DWC G2V facilities deployment models are explained with relevant details of the respective optimization objective functions. In addition, V2V energy exchange solutions are also highlighted as complementary dynamic charging solution that supports the infrastructure G2V energy transfer model. Nevertheless, further research is required into the integration of static and dynamic, G2V and V2V EV charging solutions to ensure effective penetration of EVs into the transportation industry, maximize returns on infrastructure investments and eliminate EV owners' anxiety.

References

[1] D. Patil, M. K. McDonough, J. M. Miller, B. Fahimi, and P. T. Balsara, "Wireless power transfer for vehicular applications: Overview and challenges," *IEEE Transactions on Transportation Electrification*, vol. 4, no. 1, pp. 3–37, March 2018.

[2] E. A. ElGhanam, M. S. Hassan, and A. Osman, "Design and finite element modeling of the inductive link in wireless electric vehicle charging systems," in *2020 IEEE Transportation Electrification Conference Expo (ITEC)*, 2020, pp. 389–394.

[3] C. A. García-Vázquez, F. Llorens-Iborra, L. M. Fernández-Ramírez, H. Sánchez-Sainz, and F. Jurado, "Evaluating dynamic wireless charging of electric vehicles moving along a stretch of highway," in *2016 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM)*, June 2016, pp. 61–66.

[4] Y. D. Ko and Y. J. Jang, "The optimal system design of the online electric vehicle utilizing wireless power transmission technology," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1255–1265, 2013.

[5] S. Jeong, Y. J. Jang, and D. Kum, "Economic analysis of the dynamic charging electric vehicle," *IEEE Transactions on Power Electronics*, vol. 30, no. 11, pp. 6368–6377, 2015.

[6] Y. J. Jang, E. S. Suh, and J. W. Kim, "System architecture and mathematical models of electric transit bus system utilizing wireless power transfer technology," *IEEE Systems Journal*, vol. 10, no. 2, pp. 495–506, 2016.

[7] I. Hwang, Y. J. Jang, Y. D. Ko, and M. S. Lee, "System optimization for dynamic wireless charging electric vehicles operating in a multiple-route environment," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 6, pp. 1709–1726, 2018.

[8] Z. Pantic, "Inductive power transfer systems for charging of electric vehicles," Ph.D. dissertation, North Carolina State University, 2013.

[9] Z. Bi, "Life cycle analysis and optimization of wireless charging technology to enhance sustainability of electric and autonomous vehicle fleets," Ph.D. dissertation, University of Michigan, 2018.

[10] L. Yan, H. Shen, J. Zhao, C. Xu, F. Luo, and C. Qiu, "Catcharger: Deploying wireless charging lanes in a metropolitan road network through categorization and clustering of vehicle traffic," in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, 2017, pp. 1–9.

[11] H. Ushijima-Mwesigwa, "Optimal placement of wireless charging lanes in road networks," p. 0, 2017. [Online]. Available: http://aimsciences.org//article/id/1277eb4f-7e66-4928-80db-461df6495ec0

[12] Z. Chen, F. He, and Y. Yin, "Optimal deployment of charging lanes for electric vehicles in transportation networks," *Transportation Research Part B: Methodological*, vol. 91, pp. 344 – 365, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0191261516303319

[13] R. Riemann, D. Z. Wang, and F. Busch, "Optimal location of wireless charging facilities for electric vehicles: Flow-capturing location model with stochastic user equilibrium," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 1 – 12, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0968090X15002375

[14] M. Li, X. Wu, Z. Zhang, G. Yu, Y. Wang, and W. Ma, "A wireless charging facilities deployment problem considering optimal traffic delay and energy consumption on signalized arterial," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 12, pp. 4427–4438, 2019.

[15] Z. Khan, M. Chowdhury, S. Khan, I. Safro, and H. Ushijima-Mwesigwa, "Utility maximization framework for opportunistic wireless electric vehicle charging," 08 2017.

[16] L. Maglaras, J. Jiang, A. Maglaras, F. Topalis, and S. Moschoyiannis, "Dynamic wireless charging of electric vehicles on the move with mobile energy disseminators," *International Journal of Advanced Computer Science and Applications*, vol. 6, 06 2015.

[17] D. Kosmanos, L. A. Maglaras, M. Mavrovouniotis, S. Moschoyiannis, A. Argyriou, A. Maglaras, and H. Janicke, "Route optimization of electric vehicles based on dynamic wireless charging," *IEEE Access*, vol. 6, pp. 42 551–42 565, 2018.

[18] E. A. Elghanam, M. S. Hassan, and A. Osman, "Modeling, simulation and comparison of different ferrite layer geometries for inductive wireless electric vehicle chargers," in *Modelling, Simulation and Intelligent Computing. MoSICom 2020. Lecture Notes in Electrical Engineering*, N. Goel, S. Hasan, and V. Kalaichelvi, Eds. Singapore: Springer, 2020, vol. 659, pp. 411–418.

[19] R. Tavakoli and Z. Pantic, "Analysis, design, and demonstration of a 25-kW dynamic wireless charging system for roadway electric vehicles," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 6, no. 3, pp. 1378–1393, Sep. 2018.

[20] H. Feng, T. Cai, S. Duan, J. Zhao, X. Zhang, and C. Chen, "An lcc-compensated resonant converter optimized for robust reaction to

large coupling variation in dynamic wireless power transfer," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 10, pp. 6591–6601, 2016.

[21] R. Xiong, J. Cao, Q. Yu, H. He, and F. Sun, "Critical review on the battery state of charge estimation methods for electric vehicles," *IEEE Access*, vol. 6, pp. 1832–1843, 2018.

[22] U. S. E. P. Agency. Vehicle and fuel emissions testing: Dynamometer drive schedules. [Online]. Available: https://www.epa.gov/vehicle-and-fuel-emissions-testing/dynamometer-drive-schedules

[23] A. Lorico, J. Taiber, and T. Yanni, "Effect of inductive power technology systems on battery-electric vehicle design," in *IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society*, 2011, pp. 4563–4569.

[24] A. A. S. Mohamed, L. Zhu, A. Meintz, and E. Wood, "Planning optimization for inductively charged on-demand automated electric shuttles project at greenville, south carolina," *IEEE Transactions on Industry Applications*, vol. 56, no. 2, pp. 1010–1020, 2020.

[25] Z. Liu and Z. Song, "Robust planning of dynamic wireless charging infrastructure for battery electric buses," *Transportation Research Part C: Emerging Technologies*, vol. 83, pp. 77 – 103, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0968090X17301985

[26] M. Fuller, "Wireless charging in california: Range, recharge, and vehicle electrification," *Transportation Research Part C: Emerging Technologies*, vol. 67, pp. 343 – 356, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0968090X16000668

[27] A. Foote, O. C. Onar, S. Debnath, M. Chinthavali, B. Ozpineci, and D. E. Smith, "Optimal sizing of a dynamic wireless power transfer system for highway applications," in *2018 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2018, pp. 1–6.

[28] P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wießner, "Microscopic traffic simulation using sumo," in *The 21st IEEE International Conference on Intelligent Transportation Systems*. IEEE, 2018. [Online]. Available: https://elib.dlr.de/124092/

[29] T. Wang, B. Yang, C. Chen, and X. Guan, "Wireless charging lane deployment in urban areas considering traffic light and regional energy supply-demand balance," in *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, 2019, pp. 1–5.

Multiple Objects Tracking using Radar for Autonomous Driving

Muhamamd Ishfaq Hussain, Shoaib Azam, Farzeen Munir, Zafran Khan, and Moongu Jeon
School of Electrical Engineering and Computer Science
Gwangju Institute of Science and Technology
Gwangju, Republic of Korea
email:(ishfaqhussain, shoaibazam, farzeen.munir, mgjeon)@gist.ac.kr, zafrankhan1830@gm.gist.ac.kr

***Abstract*—Object detection and tracking are the integral elements for the perception of the spatio-temporal environment. The availability and affordability of camera and lidar as the leading sensor modalities have used for object detection and tracking in research. The usage of deep learning algorithms for the object detection and tracking using camera and lidar have illustrated the promising results, but these sensor modalities are prone to weather conditions, have sparse data and spatial resolution problems. In this work, we explore the problem of detecting distant objects and tracking using radar. For the efficacy of our proposed work, extensive experimentation in different traffic scenario are performed by using our self-driving car test-bed.**

***Index Terms*—Self-Drving Car, Radar, Obstacle Detection and Tracking, Robotics, Extended Kalman Filter**

I. Introduction

For Advanced Driving Assistance Systems (ADAS) and autonomous driving, perception of the environment is essential for robotics and autonomous driving vehicles. According to the latest technical findings by the US department of transportation National Highway Traffic and Safety Administration (NHTSA), more than 90 % of road accidents are due to human mistakes [4]. Therefore, objects detection is considered paramount for safe navigation on highways and roads within the cities. In particular, object detection in autonomous driving pertains to detecting the obstacles (i.e, 2 wheel and 4 wheel vehicles, pedestrians, and any random obstacles on-road), aimed to operate the automated systems safely. Overall, to avoid any mishap, the perception system should be strong enough to accurately detect the small distant objects on the roads.

The reliability and safety of autonomous systems are dependent on vital components fig. 2 which work hierarchically. These components fig. 2 act as backbone of the autonomous systems and improve their performance. In the first phase, different types of sensors are used by the autonomous vehicles to obtain the information (raw data) about the environment in which it is operating. Subsequently, the received raw data is processed in the perceive step, to detect and track the obstacles. The planning and actuation phase is comprised of path planing, obstacle avoidance and performing a suitable action. From start to end, the most intricate and challenging job is to perceive all the possible information from the environment. Scenes having complex dynamic environment like cross-section intersection, where different maneuvers are happening at different speeds, become tedious for safe and robust planning algorithms.

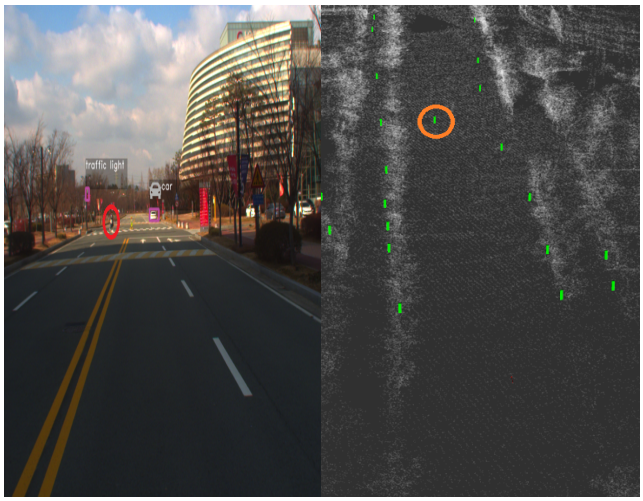


Fig. 1. The quantitative results of objects detection and tracking in real environment. Left side of the image showing the results of object detection using 'camera image' while right side displays radar's markers after removing the noise and implementing the proposed methodology for object detection and tracking.

The dawn of autonomous vehicles is emerging and will hit the public roads in coming years. The hindrance in class-specific detectors for all possible obstacles on the road may lead to complexity in training deep neural networks. In autonomous vehicles, different sensors (i.e lidar, camera, radar) are mostly used for detection and tracking the obstacles. Nowadays most frequently used sensor is lidar due to its dense and perfect range measurement. However, lidar is an expensive instrument with limited range constraints. The continuous development in deep learning vision-based detection and tracking gives us state-of-the-art results. But, training and detecting all the unknown-classes is an expensive and time taking phenomena. Besides these advancements, we are still unable to detect the distant obstacles through lidar and cameras

978-1-5386-5541-2/18/$31.00 ©2020 IEEE

TABLE I
COMPARATIVE ANALYSIS OF SENSORS IN LIEU OF OBJECT TRACKING

Sensor Modality	lidar	camera	radar
All-weather	-	✓	-
Cost	High	Low	Medium
Range	Medium (Velodyne HDL-32E)	Medium	Long Approx (200m)
Working in Darkness	-	-	✓
Direct Velocity Estimation	-	-	✓
Illumination affects	✓	✓	-

due to long-range, bad weather, and various other reasons which hamper its performance. Whereas the only leftover sensor which can accurately detect the distant obstacle is the radar. Radar bears long-range and it gives the measurement in diverse weather conditions despite of having noisy and sparse data. Radars are being widely used fro object detection and tracking due being economical, it better performance in deteriorated weather conditions and it ability to cater to the redial velocity of detected objects.

In this study, we are focusing on the importance of distant obstacles detection and tracking using radar scans data. The use of radar scans provide a promising outcome in terms of detecting the distant obstacles, and also reduce the computational cost as compared to the camera and Lidar [1]. The rest of the paper is organized as follows: Section-II explains the related work. Multi-Object Detection and tracking are discussed in section III. Section IV focuses on the experimentation and results, and finally, section-V concludes the paper.

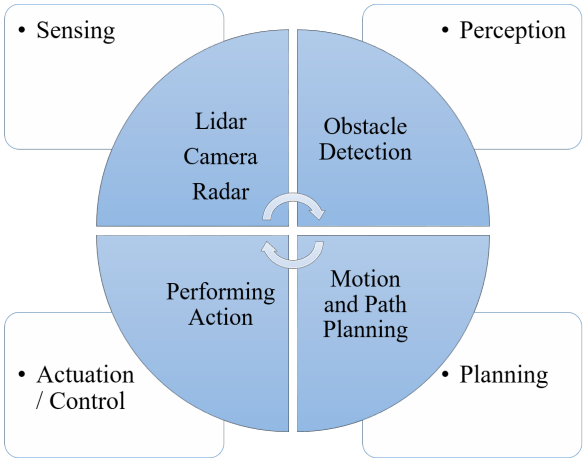


Fig. 2. The components which are showing in above figure, playing an important role in the autonomy of Autonomous Vehicle and robotics.

II. RELATED WORK

Multi-object tracking plays an important role in the advance driving assistance systems (ADAS), and also for autonomous driving. In recent years many fully convolutional detection and tracking algorithms are proposed and implemented in a real-time environment. Single-shot detectors (SSD) [7] and You Only Look Once (YOLO) [6] are using one stage algorithm which are computationally less expensive, however, they pose lower detection quality [1]. Particularly R-CNN and Faster R-CNN [11] give the state of the art solution with limited classes and higher computational cost [1]. [8] uses the radar data for obstacle detection by fusing it with a monocular camera. High-resolution radar is being used, making it possible to extract the size of the object as well as its motions precisely. This paper works on ambiguities and extended-object detection problem, which is a worth-mentioning effort in radar-based tracking [18]. Multi-modal networks are discussed which have helped us in detecting small objects at the earliest. Early fusion and segmentation are the hot research areas in tracing the distant obstacles in real-time for unmanned aerial vehicles. What, When, and how to fuse are open-ended questions in this research area which create new horizons to be explored in the world of autonomous driving vehicles [1]. Based on radar and camera [9] proposed network is a solely radar-based object detection mechanism achieving high-quality results in real-time. The same is amply related to our work as it incorporates the attention mechanism in its implementation. Exploration of resource-efficient solutions are the need of the hour for object detection in real-time. Efficiency of such solutions may be increased by allocating system resources to real objects only and ignoring distractions like pictures of vehicles on billboards.

Our work is greatly inspired by the technique used in [12], [9] and [1] which led to improved accuracy, speed, and optimization using multi-modal sensors in novel approach of attention guided multi-modality fusion mechanism [12]. This work follows a very loosely coupled technique for research in multi-modal fusion approaches.

III. MULTI-OBJECT TRACKING

Radar can directly measure the velocity of object using the doppler effect [5]. For the multiple object tracking, and to update their state dynamically, the Extended Kalman Filter (EKF) is implemented by utilizing radar measurements for different types of road entities that includes vehicle and pedestrian. Kalman filter (KF) only handles the linear data [14] while the Extended Kalman filter (EKF) [15] can handle the non-linear data as well. The measurement directly coming from the radar sensor are in raw form and carries too much noise and clutter. In order to remove the noise from the data, density-based clustering techniques are opted for raw radar data. The measurements (px, py, vx, vy) obtained from the radar scan are shown in the fig. 4 in detail. The function H

Fig. 3. Vehicle (Kia-soul EV) is used for performing experiments. The radar is mounted in the front bumper of the car. While the lidar and camera, sensors are also mounting on the roof of the car.

used for measurement is given in the Eq.1, where x is the mean of state vector, and n is the noise. The ρ is the radial distance from the origin and θ is the angle between the ρ and x. While the radial velocity is the change between the range rate $\dot{\rho}$. EKF has the built-in functionality to linearize the non-linear function using the Jacobian matrix. The filtering process includes the complete hierarchy from initialization to prediction and followed by updating the state.

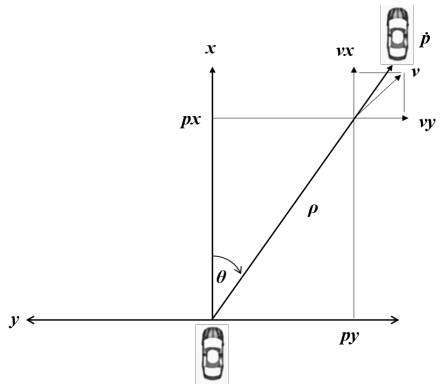


Fig. 4. The radar is giving us the range and angle of the next vehicle in the above mentioned image. And on the basis of these values, additional measurement of different components (px, py, vx, vy) is calculated, which are used in Extended Kalman Filters (EKF) for object level detection and then further used for the tracking of objects.

$$H = h(x^{'}) + n \tag{1}$$

IV. Experimentation and Results

A series of experiments were performed to ascertain the performance and quality of distant objects tracking while using radar data. The radar is mounted on the front bumper of our autonomous vehicle (KIA Soul-Ev) as shown in fig. 3. The mounting height above from the ground is $50cm$. Radar having two separate beams long and short ($\pm 45\,\text{deg}$ and $\pm 20\,\text{deg}$) with $76GHz$ pulse-doppler, while the scan interval is $20Hz$. We collected the two dimensional (2D) data of radar scans (detection). The vehicle also carried a lidar (Velodyne HDL-32E) and a monocular camera (FLIR BlackFly S), for estimating the ground truth. We performed multiple experiments to measure the accuracy and reliability of radar tracking. The experiments were performed on the main road with different environmental conditions. In order to evaluate the proposed method, we performed two different experiments. In fig. 1 left image shows the camera detection using Yolo3 [16], while the right image depicts the radar tracking data. In this image, it is clear that the camera image is unable to detect the pedestrian crossing the road, while the radar is able to detect perfectly. In experiment-II with a different scenario the camera is again unable to detect the vehicle while radar detects and tracks the same vehicle. In experiment-III, the same problem was observed during a pedestrian detection which was not detected by camera while the radar detected that person perfectly.

We created the ground truth manually, and calculated the Intersection over Union (IoU) score, to undertake the quantitative evaluation. The ground truth is based on the image data from the camera mounted on the vehicle. The qualitative results of three different experiments, are shown in fig. 1, 5, and 6 respectively. We are only focusing on the obstacles which are unable to be detected by the camera and lidar.

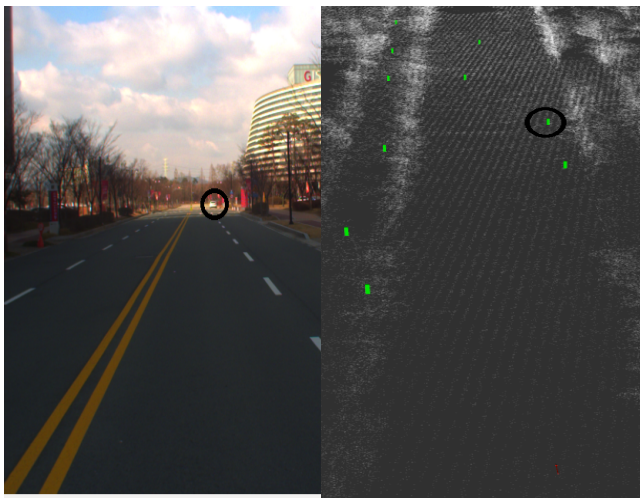


Fig. 5. In part-a showing the camera image while in part-b is the radar data of detection and tracking on real environment.

TABLE II
Intersection Over Union (IoU)

Experiment	IoU (Score)
I	80.44 %
II	79.11 %
II	80.29 %

V. Conclusion

In this work, we have evaluated the distant obstacles which are unable to be detected by the camera and lidar sensor. The same can be detected by the radar sensor by employing the

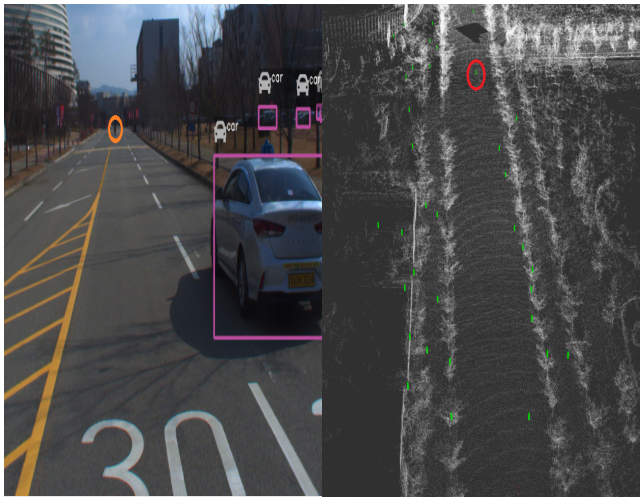


Fig. 6. In part-a showing the camera image while in part-b is the radar data of detection and tracking on real environment.

EKF. Tracking the on-road distant vehicles and pedestrians improves the safety of the autonomous vehicle's manifold despite the fact that object classification, can't be achieved with the radar data only. Detecting and then tracking such obstacles that are distracting can be further utilized in image and lidar frame to classify and pay more attention to only those particular areas which are not occupied. It is also pertinent to mention that, using radar in an autonomous stack is an increment for the safety, and reliability of autonomous driving. For the future, we will work on radar data to increase the attention of some specific regions which are neglected, due to bad weather and distance in the image or lidar frame to enhance the quality.

ACKNOWLEDGMENT

This work was partly supported by the ICT R&D program of MSIP/IITP. (2014-0-00077, Development of global multi-target tracking and event prediction techniques based on real-time large-scale video analysis) and GIST Autonomous Vehicle project.

REFERENCES

[1] S. Chadwick,w. Maddetn,and P. Newman. "Distant vehicle detection using radar and vision. In 2019 International Conference on Robotics and Automation" (ICRA) (pp. 8311-8317).

[2] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in Conference on Computer Vision and Pattern Recognition (CVPR), 2012.

[3] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[4] S. Singh. "Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey." (Traffic Safety Facts Crash•Stats. Report No. DOT HS 812 506). Washington, DC: National Highway Traffic Safety Administration. (2018, March)

[5] L. Stanislas and T. Peynot, "Characterisation of the delphi electronically scanning radar for robotics applications," 2015.

[6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp.779–788.

[7] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in European conference on computer vision. Springer, 2016, pp. 21–37.

[8] John, V., and Mita, S. "RVNet: deep sensor fusion of monocular camera and radar for image-based obstacle detection in challenging environments" In Pacific-Rim Symposium on Image and Video Technology (pp. 351-364). Springer, Cham. 2019.

[9] Nabati, Ramin, and Hairong Qi. "RRPN: radar Region Proposal Network for Object Detection in Autonomous Vehicles." In 2019 IEEE International Conference on Image Processing (ICIP), pp. 3093-3097. IEEE, 2019.

[10] Farzeen Munir, Shoaib Azam, Ishfaq Hussain, Ahmad Muqeem Sheri, and Moongu Jeon, "Autonomous Vehicle: The Architectural Aspect of Self Driving Car", Sensors, Signal and Image Processing(SSIP), 2018, Prague, Czech.

[11] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards realtime object detection with region proposal networks," in Advances in neural information processing systems, 2015, pp. 91–99.

[12] Zhang, Wenwei, Hui Zhou, Shuyang Sun, Zhe Wang, Jianping Shi, and Chen Change Loy. "Robust multi-modality multi-object tracking." In Proceedings of the IEEE International Conference on Computer Vision, pp. 2365-2374. 2019.

[13] Kim, Du Yong, and Moongu Jeon. "Data fusion of radar and image measurements for multi-object tracking via Kalman filtering." Information Sciences 278 (2014): 641-652.

[14] Welch, Greg, and Gary Bishop. "An introduction to the Kalman filter." (1995).

[15] Bolognani, Silverio, Luca Tubiana, and Mauro Zigliotto. "Extended Kalman filter tuning in sensorless PMSM drives." IEEE Transactions on Industry Applications 39, no. 6 (2003): 1741-1747.

[16] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).

[17] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2980–2988.

[18] Scheel, Alexander, and Klaus Dietmayer. "Tracking multiple vehicles using a variational radar model." IEEE Transactions on Intelligent Transportation Systems 20, no. 10 (2018): 3721-3736.

Assessing Efficiency and Aging of Lithium-Ion Battery in a Hybrid Energy Storage System

Hazem M. Sharf, Eiman A. ElGhanam, Mohamed S. Hassan, and Ahmed H. Osman
Department of Electrical Engineering
American University of Sharjah
P.O. Box 26666, Sharjah, UAE
emails: {b00062852, eelghanam, mshassan, aosmanahmed}@aus.edu

***Abstract*—This paper aims to assess the enhancements of the electric vehicle (EV) energy storage systems by hybridizing Lithium ion (Li-ion) battery with a supercapacitor, based on their operation efficiency and the aging of the overall storage system. The hybrid energy storage system brings desired improvements due to the high power density and small internal resistance of supercapacitors. However, the energy density of a supercapacitor is lower than that of a typical EV battery, making the combination of the two storage systems a promising breakthrough to increase both EV mileage and battery lifetime. Accordingly, MATLAB/Simulink tools are used in this work to set-up simulation scenarios to investigate the operation of the proposed hybrid storage system.**

I. Introduction

The race towards greener environment mandates automobile industries to work their way after improving the performance of electric vehicles (EVs). A crucial objective that manufacturers and consumers are seeking is increasing the lifetime of the EV battery to ensure sustained operation and reduce EV operational costs [1]. Accordingly, hybrid EV energy storage systems are currently being investigated, in which a supercapacitor is utilized as an auxiliary energy source to boost the battery's performance. Supercapacitors are electrolytic capacitors whose manufacturing process is different from that of ordinary capacitors, giving them a set of distinguishing features, including: high power density and fast response to electrostatic charge/discharge. Hence, these features enable the supercapacitor to be utilised for enhancing the energy storage system in electric vehicles.

While supercapacitors offer higher energy densities compared to ordinary capacitors, the energy density of a supercapacitor is significantly lower than that of the Lithium ion (Li-ion) rechargeable battery used in majority of EVs. Li-ion batteries are widely used due to their design flexibility and high energy density, as well as the desirable electrochemical properties of the Li-ion [2]. Nonetheless, Li-ion batteries exhibit low power densities [3], which degrades their efficiency in times of high current demand. This leads to high thermal stress and negatively impacts the lifetime of the battery. Accordingly, integrating a supercapacitor into the EV energy storage system together with the EV battery addresses this problem. This is because, the high power density of the supercapacitor implies that in times of high current demand, this current can be supplied by the supercapacitor instead of the battery, which thereby eliminates thermal stress and prolongs the battery's life. Moreover, by allowing supercapacitors to handle high current demands, the amount of current drawn from the battery is reduced and the vehicle's driving range is increased since the battery's state of charge (SOC) does not experience a dramatic decrease during battery discharge. The high power density of a super capacitor also enables it to charge rapidly, in comparison with the charging time of a typical EV battery. This is particularly useful in dynamic wireless EV charging systems [4], [5], in which the EV is required to receive energy wirelessly on-the-move and hence, a fast charging response time is required.

Given the aforementioned benefits of better driving range, longer battery lifetime, and less thermal stress on the battery, it is essential to develop a new energy storage system that integrates both the battery and the supercapacitor to leverage on the advantages of both and overcome their drawbacks. The challenge here pertains in the control strategy to be adopted for optimal usage of both energy sources. One of the solutions proposed in the literature involves a metaheuristic approach [6], which short-term and long-term operation plans are assigned for a given time interval and then executed by a controller. While the reported results show notable improvements, slow response time is reported to be taken by the controller to evaluate the objective function and settle on a plan. Another solution is proposed in [7], in which a controller is fed with instructions to execute based on the SOC of both the battery and the supercapacitor, which has demonstrated slight improvements in the driving range of electric vehicle. The authors in [8] suggested a new power flow energy management strategy using a model predictive control (MPC) method. This control method aims to track a set speed profile and smooth the power flowing in/out of the battery making use of the high power density characteristics of the supercapacitors.

Another difficulty faced by researchers is the optimal connection of the two energy sources. The work in [9] compares between two different connections; whether the battery is connected directly to the DC bus and supercapacitor connected to DC bus through DC/DC converters or vice versa. In their

978-1-7281-9615-2/20/$31.00

assessment, the authors concluded that having supercapacitor connected to the DC link through a DC/DC converter is better as it allows large peak currents to be provided by the supercapacitor, which is the main purpose of this electrical component. The authors in [7] assert that the optimal connection would be to have both sources connected to DC bus through a DC/DC converter to have full control over the energy flow. More issues are also faced in the design of the DC/DC converters. The work in [10] compares the conventional bi-directional topology to two improved topologies namely: three-level converter and half-controlled converter. On the other hand, paper [11] states that the usage of DC/DC converters in this application must be avoided as it produces instability to the system, and suggests a dynamical limitation method to replace it.

Accordingly, based on the aforementioned literature review, this work utilizes a simplified battery-supercapacitor interconnection topology to illustrate the enhancements that takes place in Li-ion battery when a supercapacitor is used compared to the battery functioning alone without a supercapacitor. The criteria focused on are the battery aging and the charging of a battery when wireless power transfer mechanism is used. The rest of this paper is organized as follows: Section II describes the battery and super capacitor equivalent circuit model, and presents an outline of the hybrid battery-supercapacitor EV energy storage system used in this work. Section III then provides the details of the design specifications and the simulation setup required to study the performance of the proposed hybrid EV energy storage system. The simulation results are then presented and discussed in Section III before the paper is concluded in Section V.

II. Battery-Supercapacitor Modeling

In order to understand the operation of the supercapacitor, an electrical equivalent circuit model is required. Several equivalent circuit representations are proposed in the literature [12]–[18], each of which is modeled based on different perspectives. The electrical model of the supercapacitor is provided in [12] consists of an N^th order ladder of RC blocks, which can be further simplified into a first order model, which is further expanded in [13] and [14] to study the static and dynamic behavior of supercapacitors. Another modelling procedure is utilized in [15] and [16], in which the supercapacitors are modeled to illustrate the time domain response, i.e. depending on the time response of the electrical components, which are resistors and capacitors. On the other hand, modeling based on electrochemical and thermal equations to describe the inner behavior of the supercapacitor is proposed in [17] and [18].

For this work, the first-order supercapacitor model from [12] is used. This consists of an internal equivalent series resistance, R_{C0}, a linear capacitor C_0 and a voltage-dependent capacitor, $C(V_c)$, where V_c is the terminal voltage of the supercapacitor. This is shown in Fig. 1.

On the other hand, a typical Li-ion EV battery can be modelled using a first-order electrical equivalent model consisting of a constant open-circuit voltage, V_o, connected to an ohmic

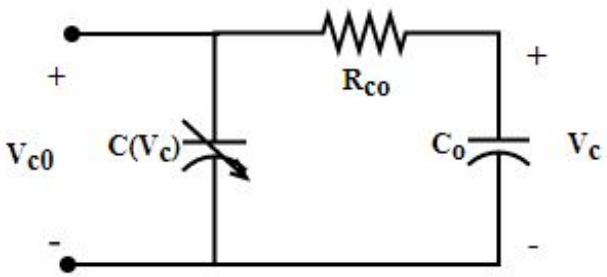


Fig. 1: First-order supercapacitor model.

resistor R_o in series with an RC network, $R_1//C_1$ that models the transient response of the battery [19]. This is shown in Fig.2.

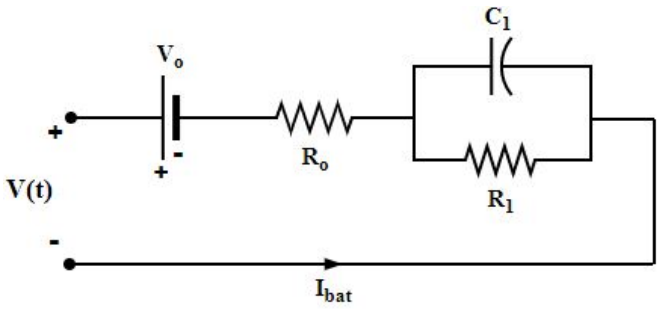


Fig. 2: First order Li-ion battery model.

Accordingly, the battery and supercapacitor can be interconnected into the hybrid battery-supercapacitor energy storage system shown in Fig. 3.

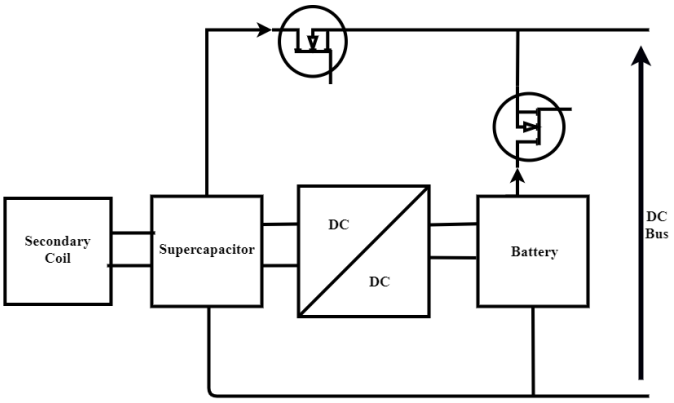


Fig. 3: Circuit schematic of the utilized hybrid energy storage system.

In this configuration, both battery and supercapacitor are connected to DC bus through N-channel MOSFETs. The switching of these transistors determine the energy management strategy for optimal share of power from both sources. Triggering $S1$ will power up the motor via battery. Motor is

powered up by supercapacitor when $S2$ is closed. a bidirectional DC/DC converter shown in Fig. 4 connects both energy sources together. This allows energy transfer between battery and supercapacitor, which is important for the supercapacitor to be recharged by the battery. Lastly, the secondary coil of the wireless charging system is connected to supercapacitor bank.

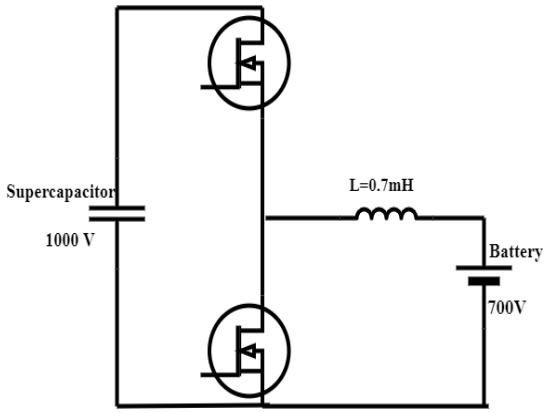


Fig. 4: Half bridge non-isolated bidirectional DC-DC converter linking supercapacitor and battery.

The topology in Fig 3 is one of several topologies proposed in the literature for the battery-supercapacitor connectivity. However, the main objective of this work is not to study the different topologies, but rather to analyze both the charging/discharging capability and battery aging performance of an EV energy storage system when integrating an supercapacitor. In particular, the energy storage system is analyzed assuming dynamic wireless power transfer is utilized for EV charging. This involves a wireless transfer of power from a primary coil laid on the road to a secondary coil fitted at the bottom of the EV, during the EV's motion [4].

Prolonging the lifetime of EV's battery is a crucial motive for equipping the energy storage system with supercapacitor bank. For that purpose, it is essential to assess the degradation of battery's lifetime over time. According to [20], lifetime degradation of a battery is measured by the change in its capacity in Ampere-hours (Ah). This can be computed using the following equation [20]:

$$\frac{1}{E_{bat,nominal}} \frac{dE_{bat,av}}{dt} = \left| \frac{dSOC}{dt} \right| a_{powercycling}, \quad (1)$$

where $E_{bat,nominal}$ represents the nominal battery capacity before any degradation, $E_{bat,av}$ is the available or useful battery capacity, SOC is the battery state of charge in percentage and $a_{powercycling}$ denotes the degradation constant. This expression can then be used to evaluate the difference in the available battery capacity based on SoC degradation with and without the use of a supercapacitor in the EV energy storage system.

III. Simulation Set-Up

The circuit schematic in Fig. 4 is implemented on MATLAB/Simulink platform to evaluate the battery charging efficiency during wireless power transfer as well as the battery's aging performance at a large number of discharge cycles, when a supercapacitor is integrated into the energy storage module. The Simulink schematic is shown in Fig. 5.

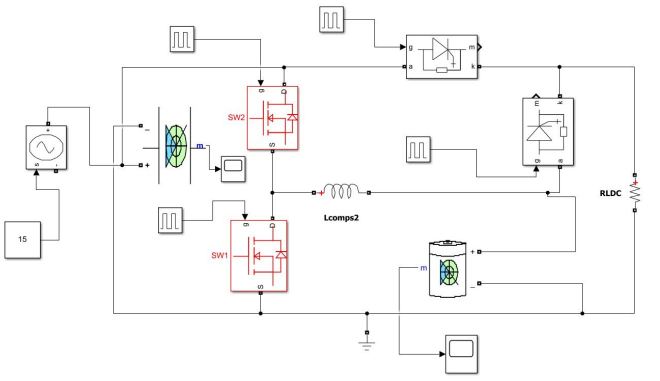


Fig. 5: Circuit schematic on Simulink.

Supercapacitor bank sizing and rating is determined by taking into considerations many factors such as: vehicle mass, air drag coefficient, motor voltage rating and power capacity. The authors in [21] explains in details the computations required for accurate sizing of supercapacitor bank. Accordingly, the specifications of the supercapacitor bank used in this work are listed in Table I.

TABLE I: Supercapacitor bank specifications.

Parameter	Value
Rated voltage	1000 V
Rated capacitance	42.85 F
Equivalent DC series resistance	8.9 mΩ
Initial voltage	0 V

Similarly, [21] lays down the design equations for battery bank design. battery bank specification are listed Table II.

TABLE II: EV battery specifications.

Parameter	Value
Rated voltage	700 V
Rated capacitance	56.3 Ah
Initial SoC	50%
Battery response time	20 s

Accordingly, two simulation scenarios are set up. The first aims to evaluate the improvement in EV battery charging efficiency when integrated with a supercapacitor, whereas the second studies the aging behavior of the EV battery with and without a supercapacitor. These are further detailed as follows.

A. Charging Efficiency

The main objective of this analysis is to investigate whether involving a supercapacitor in charging process enhances the charging capability of the EV battery. This is done by carrying out a comparison between the battery charging patterns in two different connections. Firstly, the battery is charged directly

from the secondary coil as shown in Fig. 6 and the state of charge is tracked.

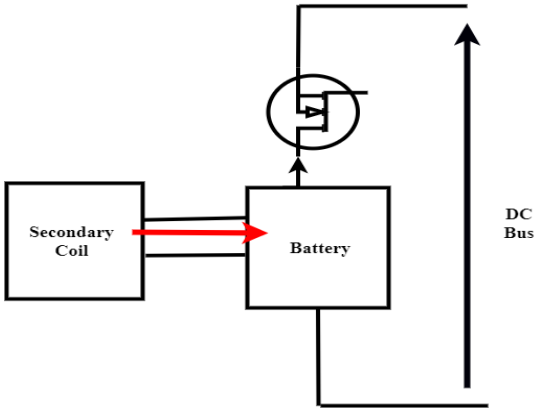


Fig. 6: Charging EV battery directly from secondary coil.

Secondly, the battery is charged through the supercapacitor, as shown in Fig. 7.

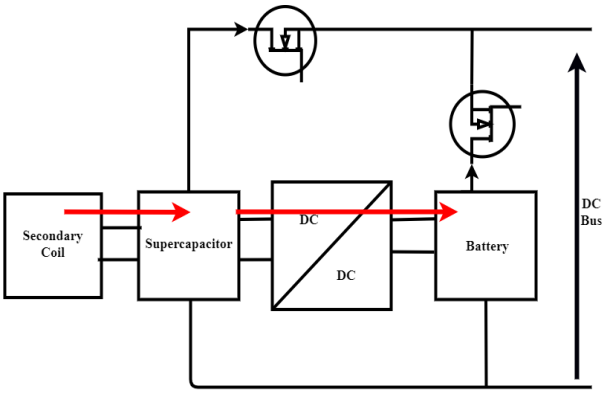


Fig. 7: Charging battery through supercapacitor.

B. Aging

The two circuits in Fig. 6 and Fig. 7 can also be used to evaluate the aging behavior of the EV battery. Nevertheless, for aging simulations, the gate signals are programmed to create a scenario in which the battery is being charged and discharged over a course of $10,000$ seconds. In one case, the supercapacitor aids the battery in providing power to the load, whereas in the other case, the battery is the sole energy source. Hence, the purpose of supercapacitor in this simulation is only to supply power to the load to ensure fair comparison, i.e. aid in power discharge to the load, but does not contribute to battery charging as in the previous simulation. The power flow diagrams are shown in Fig. 8 and Fig. 9 for illustration, with blue lines indicating power flows in and blue lines indication power flowing out.

IV. Results and Discussion

Upon simulating the circuit topologies in Section III, the following results are observed.

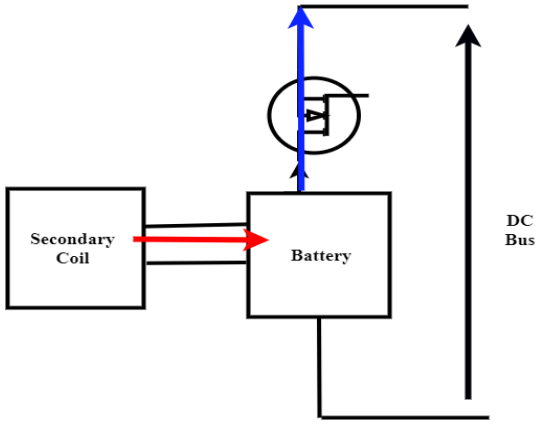


Fig. 8: Power flow without supercapacitor.

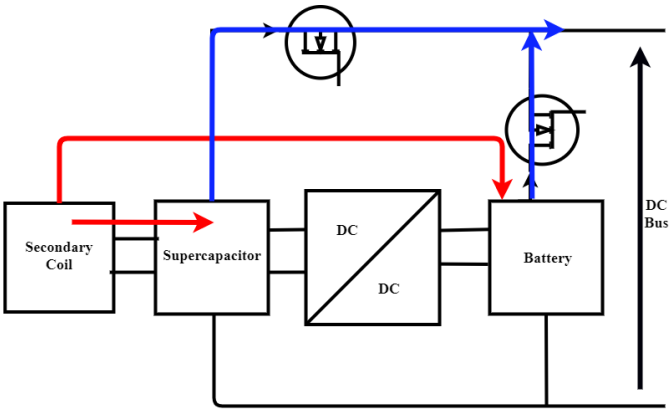


Fig. 9: Power flow with supercapacitor.

A. Charging Efficiency

As highlighted earlier, this simulation is intended to track the charging patterns for two different circuit connections. The simulated scenario entails that the vehicle will undergo wireless charging for 5 seconds, which is equivalent to turning on the power supply for 5 seconds on Simulink. The simulation reveals propitious results as shown in Fig. 10.

As observed in Fig. 10, when the battery charges directly from the coil, it does so only for an instant of time. The charging takes place as long as the two coils are coupled, and stops when no coupling between the coils occurs. On the other hand, when a supercapacitor bank is involved in the charging circuit, the charging characteristics are improved. This is because the supercapacitor continues to charge the battery after the 5 seconds charging period, due to its inherent fast electrostatic charging. Accordingly, 5 seconds are sufficient for the supercapacitor to charge enough to boost the SoC of the EV battery from 50% to 60% as compared to the minor increase from 50% to 51% when no supercapacitor is involved. In addition, by observing the simulation results, it is noted that the 9% increase in SoC took place over a duration of 15 minutes, which is significantly longer than the 5 seconds charging duration. This means that integrating a supercapacitor

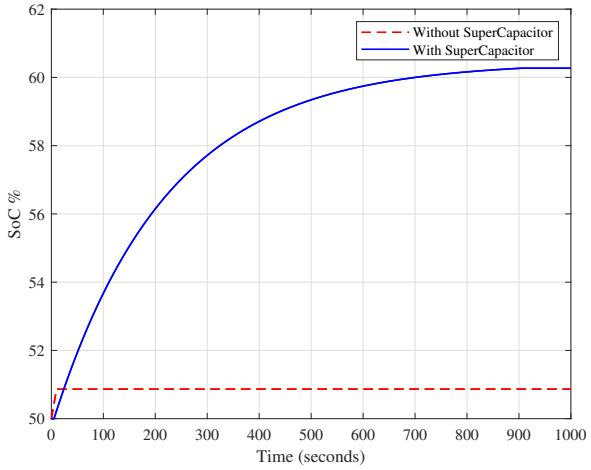


Fig. 10: State of charge vs time graph to demonstrate charging efficiency enhancement.

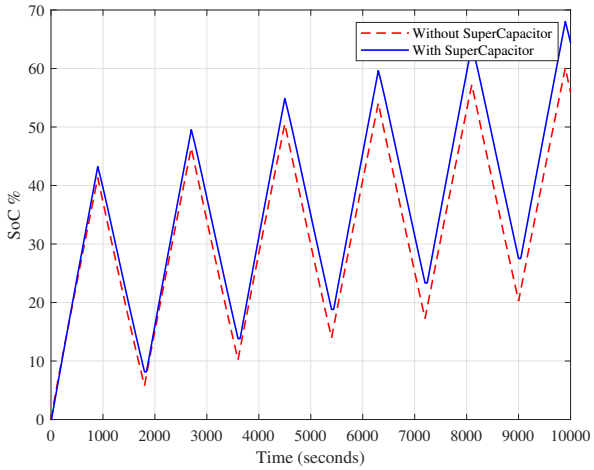


Fig. 11: State of charge vs time graph to demonstrate aging.

bank can significantly boost the EV battery charging efficiency during dynamic wireless charging by extending the effective charging duration using the stored energy in the supercapacitor. This can be further improved by having a current control loop, to control the current output from the supercapacitor and steepen the charging curve. Additionally, using different circuit topology may decrease the power loss between the two energy banks, and result in even faster charging behavior.

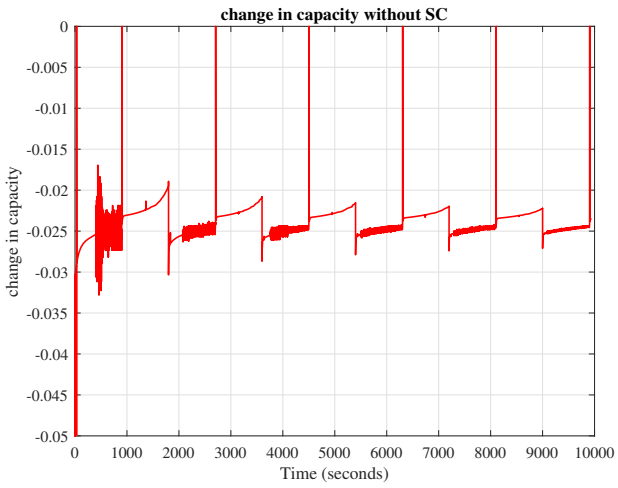
B. Aging

The second simulation measures the degradation of the Li-ion battery over the course of approximately 3 hours, using the two topologies shown in Section III-B. It should be noted that the degradation constant is picked to be larger than usual to show a noticeable amount of battery capacity degradation, because in practice, the degradation over 3 hours would be negligible.

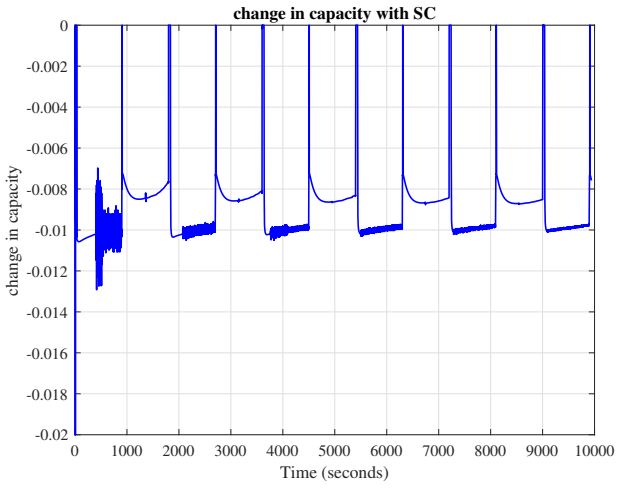
This simulation scenario is designed to allow the battery banks in the two circuit topologies to charge and discharge for the same time interval. To investigate the role of supercapacitor in this simulation and its direct impact on aging, the supercapacitor bank delivers power to the load along with battery bank and does not participate in the battery charging process. Variations in the state of charge are shown in Fig. 11. Battery charging rate in both cases are equal, but the discharge rate is slower when the supercapacitor bank is equipped in the energy storage system, in comparison with having a sole battery supplying the load. The slower discharge rate can be observed as the triangular wave starts to offset, since not all of the power is flown out from the battery.

To be able to visualize battery degradation in both topologies, equation (1) is used. The change in battery capacity over time is shown in Fig. 12.

It can be observed from Fig. 12 that the rate of change of SOC is slower when using a supercapacitor in comparison with the case when the supercapacitor is not involved. This



(a)



(b)

Fig. 12: Rate of change of battery's degradation: (a) with supercapacitor, (b) without supercapacitor.

implies slower degradation in EV battery's available capacity when a supercapacitor is integrated into the energy storage system. This is further clarified in Fig. 13 that shows the overall degradation over time. As shown in Fig. 13, Li-ion

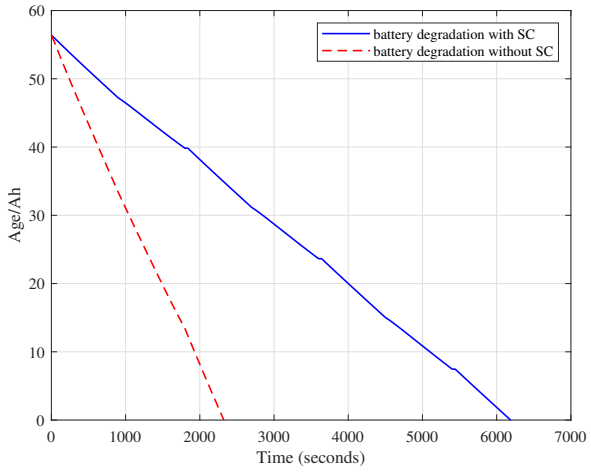


Fig. 13: Degradation of EV Battery over time.

battery lifetime increases with a hybrid energy storage system, although the degradation appears to be quite fast due to the large degradation constant used in the simulations.

V. Conclusion

In essence, a hybrid energy storage system introduces a new era of electric vehicle ubiquity, by integrating a supercapacitor in the EV energy storage module to improve the EV battery charging efficiency and aging performance. Although the results reported in this work are obtained using a simplified circuit topology, remarkable improvements are observed which can be further realised with an extensive control system and a more sophisticated energy system model. In fact, ongoing research in this field aims to find the most optimum approach for energy sharing among the two energy sources. Nevertheless, it is believed that using a supercapacitor bank as a part of the energy management system offers an extended EV driving range, and an improved EV battery lifetime, thereby addressing some of the most critical concerns hindering widespread EV penetration into the transportation market.

Acknowledgment

This work is jointly supported by the American University of Sharjah through SCRI Grant No. SCRI 18-CEN-10 and by Sharjah Research Academy.

References

[1] Fanning Jin, Mengqi Wang, and Changjian Hu, "A fuzzy logic based power management strategy for hybrid energy storage system in hybrid electric vehicles considering battery degradation," in *2016 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2016, pp. 1–7.

[2] J. M. Tarascon and M. Armand , "Issues and challenges facing rechargeable lithium batteries," vol. 414, pp. 359–367, 2016.

[3] J. M. Miller, T. Bohn, T. J. Dougherty, and U. Deshpande, "Why hybridization of energy storage is essential for future hybrid, plug-in and battery electric vehicles," in *2009 IEEE Energy Conversion Congress and Exposition*, 2009, pp. 2614–2620.

[4] E. A. ElGhanam, M. S. Hassan, and A. Osman, "Design and finite element modeling of the inductive link in wireless electric vehicle charging systems," in *2020 IEEE Transportation Electrification Conference Expo (ITEC)*, 2020, pp. 389–394.

[5] C. A. García-Vázquez, F. Llorens-Iborra, L. M. Fernández-Ramírez, H. Sánchez-Sainz, and F. Jurado, "Evaluating dynamic wireless charging of electric vehicles moving along a stretch of highway," in *2016 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM)*, June 2016, pp. 61–66.

[6] F. Machado, J. P. F. Trovão, and C. H. Antunes, "Effectiveness of supercapacitors in pure electric vehicles using a hybrid metaheuristic approach," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 1, pp. 29–36, 2016.

[7] I. Azizi and H. Rajeai, "A new strategy for battery and supercapacitor energy management for an urban electric vehicle," *Electrical Engineering*, 04 2017.

[8] L. Zhang, X. Xia, and F. Barzegar, "Control of a battery/supercapacitor hybrid energy storage system for electric vehicles," in *2017 36th Chinese Control Conference (CCC)*, 2017, pp. 9560–9565.

[9] X. Liu, Q. Zhang, and C. Zhu, "Design of battery and ultracapacitor multiple energy storage in hybrid electric vehicle," in *2009 IEEE Vehicle Power and Propulsion Conference*, 2009, pp. 1395–1398.

[10] H. Xiaoliang, H. Yoichi, and H. Tosiyoki, "Bidirectional power flow control for battery super capacitor hybrid energy system for electric vehicles with in-wheel motors," in *2014 16th International Power Electronics and Motion Control Conference and Exposition*, 2014, pp. 1078–1083.

[11] B. Nguyen, R. German, J. P. Trovao, and A. Bouscayrol, "Improved voltage limitation method of supercapacitors in electric vehicle applications," in *2016 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2016, pp. 1–6.

[12] P. J. Grbovic, P. Delarue, P. Le Moigne, and P. Bartholomeus, "Modeling and control of the ultracapacitor-based regenerative controlled electric drives," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 8, pp. 3471–3484, 2011.

[13] A. S. Weddell, G. V. Merrett, T. J. Kazmierski, and B. M. Al-Hashimi, "Accurate supercapacitor modeling for energy harvesting wireless sensor nodes," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 58, no. 12, pp. 911–915, 2011.

[14] V. Musolino, L. Piegari, and E. Tironi, "New full-frequency-range supercapacitor model with easy identification procedure," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 1, pp. 112–120, 2013.

[15] N. Bertrand, J. Sabatier, O. Briat, and J. Vinassa, "Embedded fractional nonlinear supercapacitor model and its parametric estimation method," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 12, pp. 3991–4000, 2010.

[16] H. Gualous, R. Gallay, A. Crausaz, and A. Berthon, "Frequency, thermal and voltage supercapacitor characterization and modeling," *Journal of Power Sources*, vol. 165, p. 928–934, 03 2007.

[17] Rengui Lu, Guoliang Wu, Rui Ma, and Chunbo Zhu, "Model based state of charge estimation method for ultra-capacitor," in *2008 IEEE Vehicle Power and Propulsion Conference*, 2008, pp. 1–5.

[18] A. Khaligh and Z. Li, "Battery, ultracapacitor, fuel cell, and hybrid energy storage systems for electric, hybrid electric, fuel cell, and plug-in hybrid electric vehicles: State of the art," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 6, pp. 2806–2814, 2010.

[19] B. Y. Liaw, G. Nagasubramanian, R. G. Jungst, and D. H. Doughty, "Modeling of lithium ion cells—a simple equivalent-circuit model approach," *Solid State Ionics*, vol. 175, no. 1, pp. 835–839, 2004, fourteenth International Conference on Solid State Ionics. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167273804006678

[20] I. Baghdadi, O. Briat, J.-Y. Deletage, J.-M. Vinassa, and P. Gyan, "Dynamic battery aging model: Representation of reversible capacity losses using first order model approach," 10 2015, pp. 1–4.

[21] R. Sadoun, N. Rizoug, P. Bartholumeus, B. Barbedette, and P. LeMoigne, "Sizing of hybrid supply (battery-supercapacitor) for electric vehicle taking into account the weight of the additional buck-boost chopper," in *2012 First International Conference on Renewable Energies and Vehicular Technology*, 2012, pp. 8–14.

Biomechanics of Bio-Fluid in the Microfluidic Channels Using Computer Simulations

Richard Kyung
CRGNJ(Choice Research Group)
Cresskill, NJ, USA
richardkyung@ChoiceRG.com

Aaron Zhao
Harvard-Westlake School
Los Angeles, LA, USA
AaronZhao@ChoiceRG.com

***Abstract*– In this paper, optimizing microfluidic technologies through a multiple channel network in a organ-on-a-chip (OOC) system which is a type of cell culture chip mechanics system were suggested. Using numerical and computer programming, this paper studied the micro-fluid flow in organ-on-a-chip microfluidic systems considering factors such as flow pattern, optimal flow rate and flow uniformity. For the purposes of this paper, micro-fluidic channels with a circular cross section were chosen due to its low fabrication complexity. MATLAB computer code was developed to investigate how the flow rate Q would change based on a variety of factors through both an iteration analysis using continuity equation and modified Bernoulli equation, and the Hardy-cross method which is an alternate iterative method. Ultimately, we found that flow rate Q had a quadratic relationship to length/diameter of the channel in both iterative methods.**

***Keywords* – Flow rate, micro-fluidic channels, Continuity Equation, Modified Bernoulli Equation, Hardy-Cross Method**

I. Introduction

The recent application of microfluidics in research involves examining cell growth and development in dynamic microenvironments [4]. The advantages of the microfluidic technology include miniaturization, parallelization, and integration. This research quantitatively studies the micro-fluid flow in organ-on-a-chip (OOC) microfluidic systems.

Defining necessary specifications builds the foundation in designing the structure of microfluidic channels. During this process, factors such as sample size, flow pattern, optimal flow rate, flow uniformity, and fabrication complexity must be considered.

Similar to a tubular network of blood vessels, a microfluidic channel that separates the flow into multiple channels is proposed. The goal of this particular system is to address the constraints on the microfluidic system.

Micro-fabrication provides a significant toolbox which helps process and manufacture the microfluidic channel. However, microfluidic systems are often costly due to the complexities that result from inconsistent channel depths and dimensions. This necessitates using sophisticated technology and establishing fabrication procedures. In order to fabricate the microfluidic channel, the proposed geometry must include and develop uniform cell deposition in proper devices.

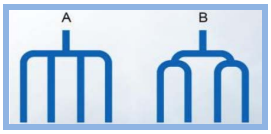
The design of the proposed microfluidic channels has circular cross sections because of its low fabrication complexity and simple manufacturing procedure. With a Reynolds number less than 2000 the laminar flow requirement of such microfluidic channels is satisfied [1].

In this study, the channel diameter D is constant at nano-scaled dimensions of 4E-9m and 3E-9m. In addition, the cross-section of the channels will be circular: a few nanometers in diameter, and a few millimeters long. Multiple channels may be used in a device by either joining them (parallel) or having many disconnected ones (independent).

II. Building Blocks of OOC Network

As in electronics, fluidic devices may be idealized to be fluidic circuits made of various independent components and elements as they are shown below [2, 4]:

(a) Multiple channels [2]



(b) Channels with parallel/wound and parallel/straight [5]

Figure 1. Schematic diagram of basic channel system

These channels can be manufactured with a variety of surface features (ridged etc.) and can be coated with various chemicals.

III. Applications Iteration Analysis of Biofluid in Micro Channel in OOC System

A. Continuity Equation and Modified Bernoulli Equation

Consider the micro channel system shown in Figure 2.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

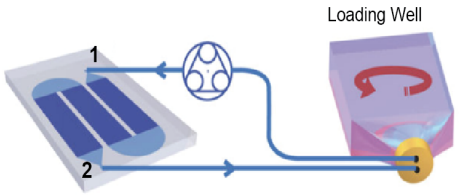


Figure 2. Schematic diagram of a channel system

The energy equation between 1 and 2 in the system is

$$\left(\frac{p}{\gamma}+\frac{V^2}{2g}+z\right)_1-\Sigma h_L=\left(\frac{p}{\gamma}+\frac{V^2}{2g}+z\right)_2 \qquad (1)$$

The head loss in the channel due to viscous or turbulent effects is given by the Darcy-Weisbach equation which is:

$$h_f=\frac{V^2}{2g}\frac{L}{D}f \qquad (2)$$

where

L = channel length (known)
D = channel diameter (known)
f = friction factor

Q is the volume flow rate and υ is the kinematic viscosity . The expression for V that was used in Equation (5) is

$$V=\frac{Q}{A}=\frac{4Q}{\pi D^2}$$

Manipulating the equations and solving for Q^2 gives

$$Q^2=\frac{g\pi^2D^4(z_2-z_1)}{8\left(\frac{L}{D}f+K_{entrance}+K_{valve}\right)}$$

Unknown value in the channel flow problem is usually the flow rate Q. Or it could be the channel diameter D. In this case an iteration scheme may be used to solve the problem. This is because velocity V is related to flow rate Q and friction factor f is related to the Re, which is related to flow rate Q. So, the last Equation is an implicit function of Q. The following iteration scheme may be used to solve for Q.

1. Assume a value for friction factor f , say f_1= 0.05. (Experiment data shows that the f ranges from 0.005 to 0.1).
2. Solve Equation (9) for Q.
3. Solve Equation (5) for Re.
4. Solve Equation (4) for friction factor f, and assign it as f_2.
5. If the difference $|f_2-f_1|<\varepsilon$ say $\varepsilon=1.0*10^{-5}$ where epsilon is tolerance, or difference between f_2 and f_1 then the flow rate Q is the correct value otherwise set $f_1=f_2$ and repeat process until condition of item 5 is satisfied.

We now create an algorithm using MATLAB program to determine the flow rate, Q, and the friction factor f by the iteration method described above.

Use the following values for the problem.

L =0.008, D=0.000004, z1=0.0001, z2=0, g=9.8 m/s^2, and nu=1.15e-06 m^2/s

Output Data

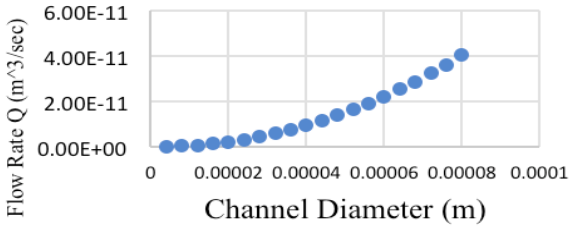


Figure 3. Channel Diameter (m) vs Flow Rate Q (m^3/sec)

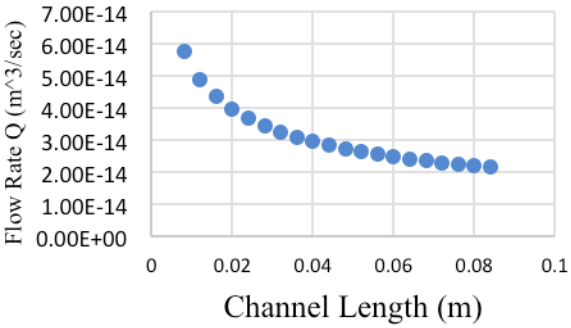


Figure 4. Channel Length (m) vs Flow Rate Q (m^3/sec)

B. The Hardy-Cross Method

Geometrical complexity of the microfluidic channel

The Hardy-Cross method is another iterative method that can be used to find flow rates in OOC network and head losses throughout a 2-D channel network. A loop is defined by a closed path and

Schematic Diagram of a 2 Dimensional Rectangular Micro-channel

The fluid flow is assumed to be laminar, steady state and incompressible with constant properties while dissipation, pressure work and body forces are neglected. In this tree-like channel, fluid flows first from the inlet to the bifurcation of the inlet, then flows further downstream to where additional bifurcation occurs. The fluid flows through three sub-

channels, which are then merged into two outlets as it is shown in the Figure 5.

The following definitions are used in describing the Hardy-Cross Method. (see Figure 6).

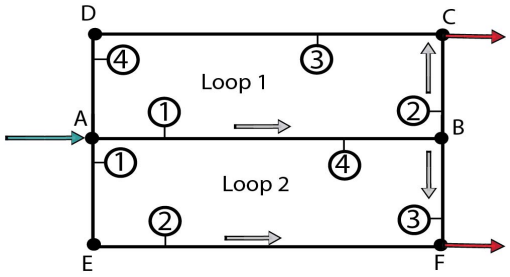


Figure 5. Loop rule - Two loops of OOC network

$$Q_{11} > 0, \quad Q_{24} < 0, \quad Q_{11} = -Q_{24}, \quad Q_{22} > 0$$

A two dimensional microfluidic system with the length of L and the width of W is considered as shown in Figure 5. The length is longer than the width of the channel (W< L).

Loop Rules

1. The flow rate Q is positive if the flow is in the counterclockwise direction around the loop.

2. The head loss h_f is considered positive if the flow is in the counterclockwise direction around the loop.

3. The two loops with a common channel can be a positive Q in one loop and a negative Q in the other loop.

Loop Number, Channel number , Channel Length (m),Channel diameter (cm) are assumed as follows :

TABLE I. CHANNEL LENGTHS AND DIAMETERS

Loop Number	Channel number	Channel Length (m)	Channel diameter(m)
1	1	0.00004	0.000000004
1	2	0.00001	0.000000004
1	3	0.00004	0.000000003
1	4	0.00001	0.000000003
2	1	0.00001	0.000000003
2	2	0.00004	0.000000004
2	3	0.00001	0.000000003
2	4	0.00004	0.000000004

Loop Number, Channel number and Initial guess Q (m3/s) are set as follows:

TABLE II. INITIAL GUESS Q FOR THE CHANNEL SYSTEM

Loop Number	Channel number	Initial guess Q (m3/s)
1	1	0.15
1	2	0.05
1	3	-0.05
1	4	-0.05
2	1	0.10
2	2	0.10
2	3	-0.10
2	4	-0.15

Output Data

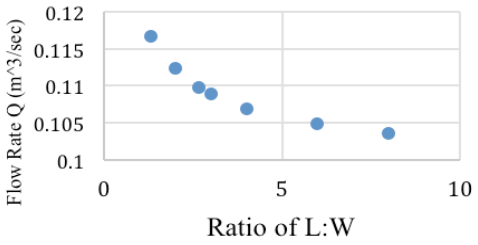


Figure 6. Ratio of L:W vs Flow Rate Q for Channel (1,1) Starting D

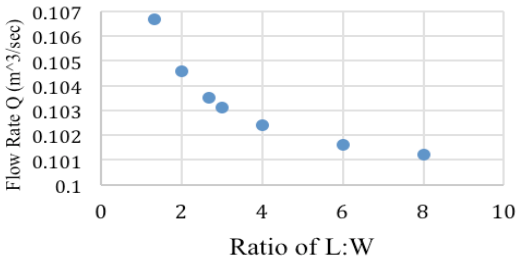


Figure 7. Ratio of L:W vs Flow Rate of Q for Channel (2,4) Invert Starting D

General Trends for (2,3): Q values increase as Ratio of L:W increases regardless of D value. As we move from a starting D to an invert starting D, the Q values in general increase by an average of 8.5E-4.The average difference in values shown here are identical to the one found in channel (2,1);(2,2) but instead of decreasing it is increasing.

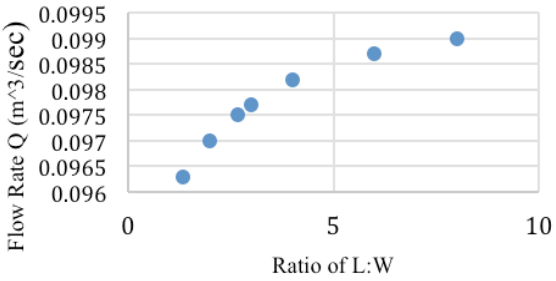


Figure 8. Ratio of L:W vs Flow Rate of Q for Channel (2,3) Invert Starting D

IV. CONCLUSION

Through two separate iterative methods that accounted for viscous and turbulent effects that lead to head loss throughout a two-dimensional channel network. Computer code was then developed to investigate how the flow rate Q would change based on a variety of factors.

First, an Iteration Analysis using the Continuity Equation and Modified Bernoulli Equation was done and found that flow rate Q (m^3/sec) exhibited quadratic growth to channel diameter(m) and quadratic decay with channel length(m).

Then, an alternate iterative method, the Hardy-Cross Method, was performed and found that flow rate Q(m^3/sec) also exhibited quadratic properties in relation to ratio of length to width of channel (m). This relationship was much more nuanced as it considered varying length to width ratios as well as varying channel diameters in the entire loop. Channels (1,1), (1,2), (2,1), (2,2) and (2,4) all saw a decrease of flow rate Q as ratio of length to width increased regardless of channel diameter values. Channels (1,3), (1,4) and (2,3) all saw an increase in flow rate Q as ratio of length to width increased regardless of channel diameter values.

Ultimately, microfluidic channels can be made more efficient using this proposed system or separating flow into multiple channels. Flow rate may be increased and decreased based on channel length, width and diameter. Further studies may may be focused on variations in channels such as uneven and ridged surfaces, chemically coated surfaces, or multi-layer photolithographically manufactured surfaces.

REFERENCES

[1]https://en.wikipedia.org/wiki/Reynolds_number

[2]http://www.princeton.edu/prism/microfluidics/training/bootcamp/2015__Materials/bootcamp1_1_for_printing.pdf

[3]https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4129156/

[4]https://quickersim.com/cfdtoolbox/tutorial/tutorial-16-laminar-flow-microfluidic-device/

[5]https://lben.epfl.ch/files/content/sites/lben/files/users/179705/Lab%20on%20a%20Chip%20Handout.pdf

Model Predictive Speed Control of DC-DC Buck Converter Driven DC-motor with Various Load Torque Values

1st Mohamed E Albira
Electrical and Computer Engineering
Oakland University
Rochester, Michigan, USA
mohamedalbira@oakland.edu

2nd Abdullah Alzahrani
Electrical and Computer Engineering
Oakland University
Rochester, Michigan, USA
alzahrani2@oakland.edu

3rd Mohamed Zohdy
Electrical and Computer Engineering.
Oakland University
Rochester, Michigan, USA
zohdyma@oakland.edu

***Abstract*__ This paper presents a speed control of a buck converter driven a direct current (DC) motor. The speed control is designed and programmed to compute a constrained model predictive controller (MPC) algorithm that regulates a reliably constant and accurate speed in the presence of different value of load torque. The challenge of regulating speed under variable load torques is solved by using the predicted discrete averaged state space model and the predicted constrained input control signal. The proposed strategy uses the output feedback signal to the MPC to first predict the state vector and the input control signal that are fed to the DC-DC buck converter to adjust the required output voltage and current, then fed to the permanent magnet DC-motor. Since the DC-motor output speed signal varies, a constrained input control signal applied to the DC-DC converter to keep the control signal limited between boundaries of 0-1. This process is fulfilled by computing the quadratic programming (QP) optimization algorithm at each prediction horizon interval. MATLAB and Simulink were used to obtain the simulated results and output figures.**

Keywords— DC-DC Buck converter, DC-motor, MPC, QP Optimization, MATLAB.

I. INTRODUCTION

DC-motor is one of the main components when it comes to the high precision motion control systems, and its accuracy made it extensively required in many industrial instruments, tools, and devices. Many applications use the DC-motor for their motion operations. For instance, small applications such as toys, and large applications such as electric vehicles, elevators, and hoists [3].

Although the DC-motor is a simple and accurate, it requires a controller to regulate the armature current, voltage, and the speed/position. One of the most commonly used controlling methods is the pulse-width modulation (PWM) technique. The PWM technique works well but the nonlinearity of the converter switching can cause a rough start of the motor and variations in the armature voltage and current, which could damage the DC-motor mechanism by placing too much stress on the electronic components of the DC-motor circuitry [2]-[4].

DC-DC buck, boost, and buck-boost converters are effective resolutions for these types of problems because they contain the energy-storage (e.g., inductor and capacitor) components which are capable of generating the required noiseless voltage and current to start the DC-motor smoothly. Combining the DC-DC converter with DC-motor also can improve the velocity tracking performance of the DC-motor by providing purified output voltage and current to the DC-motor [5] and [6]. The buck converter is the most commonly in use of the above-mentioned converters because of its function to step down the input voltage to the DC-motor [5].

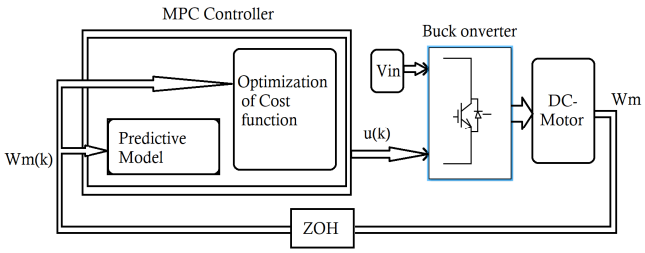


Fig. 1. The schematic diagram of DC-DC buck converter driven DC-motor with MPC controller

The advantages of using the DC-DC converters to drive a DC-motor are that they easily can achieve a smooth start and quickly can regulate the sudden change of the motor speed by reaching the desired value of the armature voltage and current [3]. However, attaining accurate and regulated speed can be challenging because of the nonlinear switching behavior of the DC-DC converters, which can cause mutation in the armature current and voltage that might lead to sluggish in the motor speed [3] and [7].

Many attempts have been made to design a control system to enhance the performance of the higher precision angular velocity/speed DC-DC converter driven DC-motor. The study [8] used two feedback control strategies, the linear quadratic regulator (LQR) and the conventional proportional–integral–derivative (PID) controllers - to track the angular velocity reference trajectory of a DC-DC buck converter driven DC-motor, but both strategies were not appropriate to track the reference trajectory. In the work [5], it has been designed an adaptive controller to control the velocity of DC motor driven by a DC-DC buck converter with a loading torque on a DC motor that was applied as an unknown. The research [3] a new discrete-time robust MPC algorithm was investigated to control the speed of a DC-DC buck converter driven DC-motor with the presence of multiple disturbances and parametric uncertainties, they used a discrete-time reduced order, generalized proportional-integral observer (GPIO) to estimate the virtual system states and the disturbances. The outcome described in [9] is an extension of the work described in [10] where a state dependent switching function with limited switching frequency has been designed and applied to the speed control of a DC motor driven by a buck-boost converter.

This research investigates an effective constrained MPC controller for speed control regulation of a DC-DC buck converter driven DC-motor based on smooth trajectory tracking as it is shown in Fig. 1. The combined discrete time-averaged plant model over the prediction horizon to predict the future attitude of the controlled system model. The quadratic programming (QP) optimization algorithm

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

subjected to unequal constraints solved to predict the input control signal (duty cycle) which govern the buck converter output voltage and current at every sampling time interval. This operation allows the MPC controller to form the input control signal to the DC-DC buck converter to regulate the outputs voltage and current for the DC-motor output speed to track the trajectory reference signal. Different speed and load torque were applied to examine the efficiency of the proposed control system.

The following section II describes a DC-DC buck converter driven DC-motor dynamic model and the averaged discrete time model design. Section III describes the MPC design. Section IV and section V present the practical design of the MPC and the performance of the proposed methodology by numerical simulation results. Finally, the conclusion is described at the end of this research.

II. DC-DC Buck Converter Driven DC-motor Dynamic Model

A. The DC-DC Buck converter driven DC-motor components and Parameters

The proposed model of the DC-DC buck converter driven DC-motor circuit shown in the Fig. 2 contains two parts the generic PWM-based DC-DC buck converter, and the permanent magnet DC-motor. The model is a linear fourth-order system that can be obtained as a linear state-space model system as in equation (1), by applying the Kirchhoff's voltage and current laws (KVL, KCL) to the circuit shown in Fig. 2. The mathematical model can be derived to illustrate the inductor current i_L, capacitor voltage vo, the armature current i_m, and the angular velocity w_m [11].

$$\begin{aligned} \overline{\boldsymbol{x}} &= A\boldsymbol{x} + Bu \\ y &= C\boldsymbol{x} + Du \end{aligned} \tag{1}$$

where the state vector $\boldsymbol{x} = [i_L, vo, i_m, w_m]^T$, u is the input variable, and y is the system output.

The first part of circuit shown in Fig. 2 describes: the parameters r_S for the internal resistance of the source, La for converter inductor, r_L for internal resistance of the inductor, Cr for the converter capacitor, the duty cycle δ, Vin for the voltage source, and Ts is the switching time period. The parameters for the second part are: L_m is the motor inductor, the R_m is the armature winding resistance, ke is the counter electromotive force constant, k_m is the motor torque constant, J_0 is the moment of inertia, b is the motor viscous friction coefficient, T_L is the load torque, and T_{fric} is the friction torque [11].

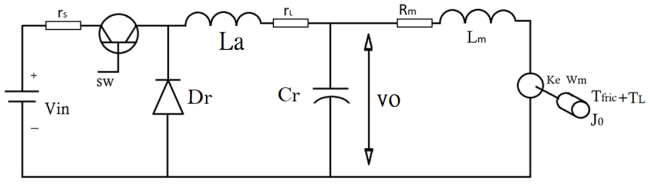


Fig. 2. The circuit diagram of the buck converter driven DC-motor

B. The mathematical model and the averaged state space model

The mathematical model can be derived to take the form of equation (1), the equivalent system is depending on the switch behaviors, this behavior generates two configurations of the system depending on the switching position. Also, this behavior directly affects the Buck converter voltage and current outputs as well as the DC-motor angular velocity by the change of the switch ON and OFF; this behavior can be expressed as an input binary signal $u = \{0,1\}$, with the consideration of the continues condition driving mode of the proposed system model, the dynamic status of the switching behavior represented to behave (ON=1 and OFF=0).

The first operation mode expressed when the switch was closed ($sw = 1$) the system represented as in (3) [11]-[13].

$$A_{on} = \begin{pmatrix} -\frac{(r_L + r_S)}{La} & -\frac{1}{La} & 0 & 0 \\ \frac{1}{Cr} & 0 & -\frac{1}{Cr} & 0 \\ 0 & \frac{1}{La} & -\frac{Ra}{La} & -\frac{ke}{La} \\ 0 & 0 & \frac{k_m}{J_0} & -\frac{b}{J_0} \end{pmatrix}, B_{on} = \begin{pmatrix} \frac{Vin}{La} \\ 0 \\ 0 \\ \frac{-(T_{fric}+T_L)}{J_0} \end{pmatrix} \tag{3}$$

The second operation mode can be expressed when the switch is opened $sw = 0$, the system represented as in equation (4) [11]-[13].

$$A_{off} = \begin{pmatrix} -\frac{r_L}{La} & -\frac{1}{La} & 0 & 0 \\ \frac{1}{Cr} & 0 & -\frac{1}{Cr} & 0 \\ 0 & \frac{1}{La} & -\frac{Ra}{La} & -\frac{ke}{La} \\ 0 & 0 & \frac{k_m}{J_0} & -\frac{b}{J_0} \end{pmatrix}, B_{off} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{-(T_{fric}+T_L)}{J_0} \end{pmatrix} \tag{4}$$

For the control aim, the averaged model can be expressed by adding the corresponding state space matrices (3) and (4) and multiplying them with the duration of one switching period (duty cycle δ) to calculate the approximate state space averaged matrices as in the equation (5) [14].

$$\begin{aligned} A_{con} &= A_{on} * \delta + A_{off} * (1 - \delta) \\ B_{con} &= B_{on} * \delta + B_{off} * (1 - \delta) \end{aligned} \tag{5}$$

By combining the equations (3) and (4) with the equation (5) and by applying the parameter's values in the Table I, the averaged state matrices can be displayed as shown in equation (6)

$$\begin{aligned} A_{con} &= \begin{pmatrix} -889.2034 & -404.3672 & 0 & 0 \\ -2.1612e4 & 0 & -2.1612e4 & 0 \\ 0 & 854.7009 & 01.4777e3 & 56.6667 \\ 0 & 0 & 576.5217 & -1.200 \end{pmatrix} \\ B_{con} &= \begin{pmatrix} 9.7257e3 \\ 0 \\ 0 \\ -333.91 \end{pmatrix} \\ C_{con} &= (0\ 0\ 0\ 1) \\ D_{con} &= 0 \end{aligned} \tag{6}$$

where the equation (6) represents the contentious time averaged state space matrices of the proposed plant model.

C. The discrete time of the averaged plant model

Discretization of the continuous averaged model is needed for designing the discrete-time MPC. The discrete mechanism that is used in this study is the zero-order hold (ZOH) method. The system model in equation (6) will be discretized over a sampling time Ts. The investigated discrete system in equation (6) will be calculated using the following expression

$$A_d = e^{A_{con}*Ts}\ ,\ B_d = \int_0^{\Delta Ts} e^{A_{con}*Ts} B_{con}\, dTs \qquad (7)$$

However, since MATLAB is used, the discretization can be done through the built-in c2d function that is commonly used for calculating the proposed discrete representation of the continuous matrices. Therefore, the discrete state equation will take the format as follows

$$\begin{aligned} z(k+1) &= A_d * z(k) + B_d * u(k) \\ y(k) &= C_d * z(k) \end{aligned} \qquad (8)$$

where the $z(k)$ is the state vector and $u(k)$ is the input variable, and the $y(k)$ is the system output at k sampling period.

III. Model Predictive Control (MPC) Design

A. Prediction Process Design

To predict the future state mode of the controlled system model, the discrete averaged model (8) will be calculated over the prediction horizon for Np intervals to predict the future position of the state vector $Z(k)$, the input control signal $U(k)$, and the output that will become $W(k)$. All will be formulated considering the following expressions (9), (10), and (11) respectively. [1], [3], [15]and [16].

$$U(k) = [u(k|k)\ u(k+1|k)\ \dots\ u(k+Np-1|k)]^T \qquad (9)$$

$$Z(k) = [\,z(k+1|k)\ z(k+2|k)\ \dots\ z(k+Np|k)]^T \qquad (10)$$

Therefore, angular velocity W can be predicted using the expression in the following equation (11)

$$W(k+1) = V * Z(k) + P * U(k) \qquad (11)$$

where

$$V = \begin{bmatrix} C_d A_d \\ C_d A_d^2 \\ \vdots \\ C_d A_d^{Np} \end{bmatrix}$$

$$P = \begin{bmatrix} C_d B_d & 0 & \cdots & 0 \\ C_d A_d B_d & C_d B_d & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ C_d A_d^{Np-1} B_d & C_d A_d^{Np-2} B_d & \cdots & C_d B_d \end{bmatrix} \qquad (12)$$

B. Optimization and Cost Function Calculation

The main goal of the control design is to force the DC-motor speed to track the angular velocity reference signal $r(k)$ along the prediction horizon. The reference signal can be defined as $r(k) = [r(k+1), r(k+2), \dots, r(k+Np)]^T$. Therefore, the equations (9) and (10) will be used to form the general quadratic cost function in given as follows

$$J(k) = \sum_{i=0}^{Np-1} \Big[\big(W(k+i|k) - r(k+1)\big)^T Q\big(W(k+i|k) - r(k+1)\big) + U^T(k+i|k) R U(k+i|k)\Big] \qquad (13)$$

with consideration of Q *and* R positive definite matrix (Q can be semi-definite). Replacing the angular velocity $W(k)$ into the expression (13) will produce the following equation (14). [1],[3],[15], and [16].

$$J(k) = U^T(k)\Omega U(k) + 2Z^T(k)\psi^T U(k) + Z^T(k)\varphi Z(k) \qquad (14)$$

where

$$\Omega = P^T \bar{Q} + \bar{R} \qquad (15)$$

$$\psi = P^T \bar{Q} V - r^T \bar{Q} P \qquad (16)$$

$$\varphi = V^T \bar{Q} V + \bar{Q} \qquad (17)$$

where

$$\bar{Q} = \begin{bmatrix} Q & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & Q & 0 \\ 0 & \cdots & 0 & Q \end{bmatrix},\ \bar{R} = \begin{bmatrix} R & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & R & 0 \\ 0 & \cdots & 0 & R \end{bmatrix} \qquad (18)$$

This study has been organized to control the angular velocity of the proposed DC-DC buck converter driven DC-motor. Since the converter used to regulate the required voltage and current fed to the DC-motor, this requires a specific amount of regulated voltage and current to achieve this purpose, generating the right input control signal $u(k)$ will directly influence the converter to regulate the desired amount of the output voltage and current for the DC-motor. To this end, applying bounded constraints will dictate the converter to provide the right outputs needed for the DC-motor to track the reference signal. The linear constraint formulas (19) and (20) applied to the input control signal are shown below, both of which are subject to the proposed cost function (14) [1],[3],[15],and [16].

$$\underline{u} \le u(k) \le \bar{u} \qquad (19)$$

which can be formulated as

$$\begin{bmatrix} I \\ -I \end{bmatrix} U(k) \le \begin{bmatrix} 1\bar{u} \\ 0\underline{u} \end{bmatrix} \qquad (20)$$

where the 1s are indicating a stacked format of ones, and the 0s are indicating a stacked format of zeros. The expression (20) will be reformulated as inequal constraints to take the form of equation (21)

$$\mathrm{M} * U(k) \le \boldsymbol{\lambda} \qquad (21)$$

The M and $\boldsymbol{\lambda}$ are inequality constraint matrices that can be calculated offline. The optimization process of the quadratic cost function (14) and the subjected constraints (21) can be solved using a QP problem [1], [15], and [16].

$$\begin{aligned} &\boldsymbol{min}\ 0.5(U^T * \Omega * U) + Z^T(k) * \psi^T * U \\ &\boldsymbol{subj\ to}\ \mathrm{M} * U \le \boldsymbol{\lambda} \end{aligned} \qquad (22)$$

Since Ω is a hessian matrix and is a positive (semi-definite) matrix, the cost function in equation (22) is a quadratic and the constraints are linear. Therefore, the generated predicted input sequence is a unique solution because it is a convex optimization problem.

Since the MATLAB is used, solving the optimization problem can be done using the built-in quadratic programming (***quadprog***) function, the expression (25) used to calculate the predicted input control signal $U_{pred}(k)$. Since the (***quadprog***) function generates a vector of elements, only the first element is needed in the process. The formula (23) will be used to take only the first element of the predicted elements in every Np sampling interval to derive the $U(k)$ as follows

$$U(k) = -[1\ 0\ 0\ \dots\dots\ 0] * U_{pred} \qquad (23)$$

IV. THE PRACTICAL DESIGN OF MPC

The proposed MPC controller's implementation requires the numerical model of the proposed DC-DC buck converter driven DC-motor. Table I describes the parameter values: the discrete time-averaged A_d and B_d matrices (7) were calculated for the prediction process. Solving the quadratic programming optimization problem has been completed using the equations (25) and (26) with consideration of the prediction horizon $Np = 10$ and the matrices Q and R that have been chosen as in equation (24), The equation (23) has been used to take the first element in every sampling instant over the prediction horizon Np. [1], [15] and [16].

$$R = 0.01, \quad Q = \begin{bmatrix} 1000 & 0 \\ 0 & 1000 \end{bmatrix}_{Np} \qquad (24)$$

the MATLAB built-in function ***quadprog*** used to solve the problem

$$U_{pred}(k) = \boldsymbol{quadprog}(\Omega, \psi, \mathrm{M}, \boldsymbol{\lambda}) \qquad (25)$$

where M and $\boldsymbol{\lambda}$ are the constraints [15], and [16].

$$\mathrm{M} = \begin{bmatrix} I \\ -I \end{bmatrix}_{Np}, \qquad \boldsymbol{\lambda} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{Np} \qquad (26)$$

Table I. The specifications of the buck converter and DC-motor:

Parameters		Values
Input voltage	Vin	$40v$
Converter Inductor	La	$2.473e-3$ H
Converter Capacitor	Cr	$46.27e-6\ F$
Switching frequency	f	$10\ khz$
Duty cycle	δ	0.6
Internal resistance of the source	r_S	$0.84\ ohm$
Internal resistance of the inductor	r_L	$0.695\ ohm$
Armature resistance	R_m	$2.729\ ohm$
Armature inductor	L_m	$1.73H$
Viscosity fraction coefficient	b	$0.000138\ N.m/rad/sec$
Inertia moment	J_0	$Kg.m^2$
Motor torque constant	k_m	$0.0663\ N.m/A$
Voltage constant	k_e	$0.0663\ V/rad/sec$
Fraction torque	T_{fric}	$0.0284N.m$
Load torque	T_L	Variable $N.m$

Table II. The steady state error and the settling time at different load torques and time intervals.

Load Torque $T_L\ N.m$	Time intervals (s)	Steady state error	settling time (ms)	Rise time (ms)
$20\ e-3$	0 - 0.5	0.15%	14.4	21.9
	0.5 - 1	0.24%	14.7	22.2
$40\ e-3$	0 - 0.5	0.25%	15.2	23.0
	0.5 - 1	0.36%	15.5	23.3

V. SIMULATION RESULTS

This section represents the configuration of the proposed control scheme which shows in Fig.1. implemented and functioned using the MATLAB and SIMULINK (version R-2017B) which are carried out in a laptop computer with CORE i7 1.8GHz CPU, RAM of 16 GB, and Windows 10 professional operating system. The controller designed by solving the quadratic programming optimization algorithm, subjected to an inequal input constraints, with control horizon $Np = 10$ time intervals, and the sampling frequency is $fs = 10\ kHz$. The DC-motor is required to track the assigned reference signal at various stages with different values of the load torque. In Fig. (3) and (5) show the angular velocity wm output and the angular velocity reference signal wm^*. It is clear that the angular velocity is tracking the reference signal at any assigned speed, even with different load torque values. The performance of the controller was barely affected by the change of the motor speed and of the various load torque values T_L. The load torque assumed with values of:

$$T_L = \begin{cases} 20e-3\ N.m \\ 40e-3\ N.m \end{cases} \qquad (27)$$

The applied speeds and load torque represent how the system reacted in terms of the steady-state error (SSE), and the settling time (St) responses. Table II. shows the numerical values for the SSE and St at each time period and each load torque value. Also, the figures (4) and (6) show that the level of the armature current i_m, inductor current i_L, and the output voltage vo. that were used as an input to the DC-motor were well-regulated to operate the system model as desired. The buck converter functioned perfectly to step down the output voltage, so it was clear that the output voltage was lower than the input voltage $V_{in} = 40v$. Although the outputs hit the highest values at the beginning and when the change of the speed occurred, the outputs were always less than the input voltage.

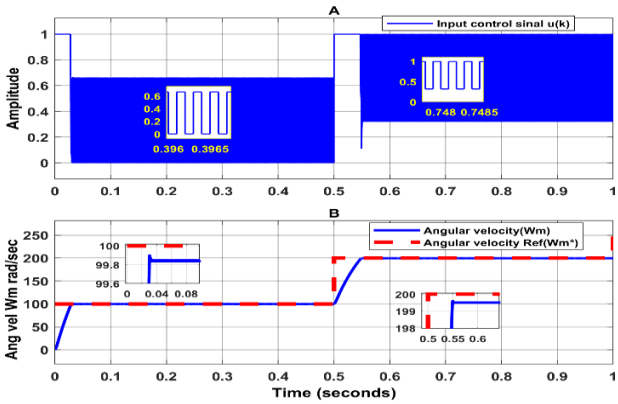


Fig 3. A. Input control signal $u(k)$, B. Angular velocity w_m, Angular velocity reference w_m^*.

At the $T_L = 20e-3\ N.m$ the Fig. 3 (A) shows the ratio of control signal that needed for the system to get to the speed of 100 rad/sec in the period between 0-0.5 s and the speed of 200 rad/sec in period 0.5-1s. The Fig. 3 (B) signifies the reference signal and the DC-motor speed. The SSE is about 0.15% to 0.24%, the Table II. shows more details of the system responses.

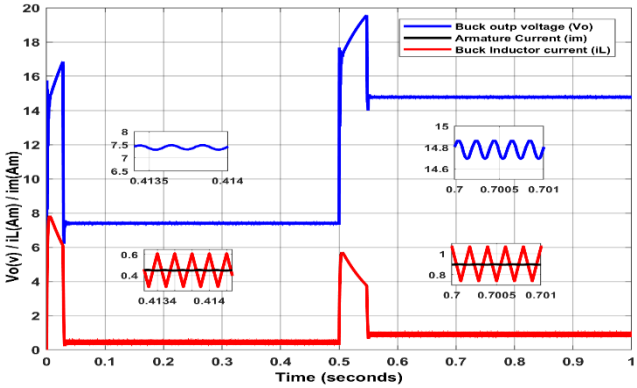


Fig 4. The outputs voltage vo, inductor current i_L, and the armature current i_m.

At the value of $T_L = 20m\ N.m$ it is clear in the Fig.4 that the vo is about $7.5v\ to\ 14.7v$, and i_m and i_L are approximately between $0.45A$ to $0.9A$.

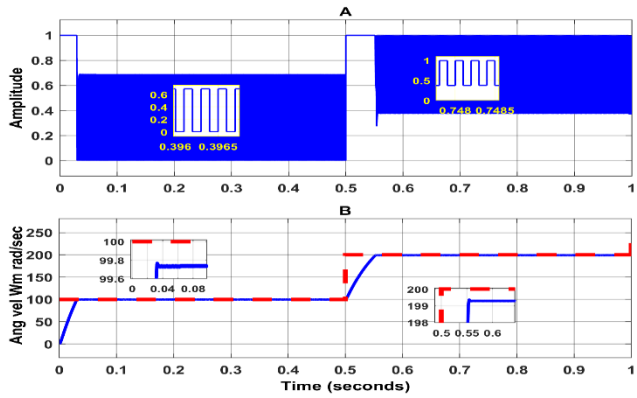


Fig 5. A. Input control signal $u(k)$, B. Angular velocity w_m, Angular velocity reference w_m^*

The Fig. 5 (A) shows the control signal needed for the buck converter to regulate the required voltage and current for the DC-motor. Fig. 5 (B) indicates that the DC-motor speed signal is tracking the reference signal. The SSE percentage is about 0.25% to 0.36%, the Table. II shows more details of the system responses.

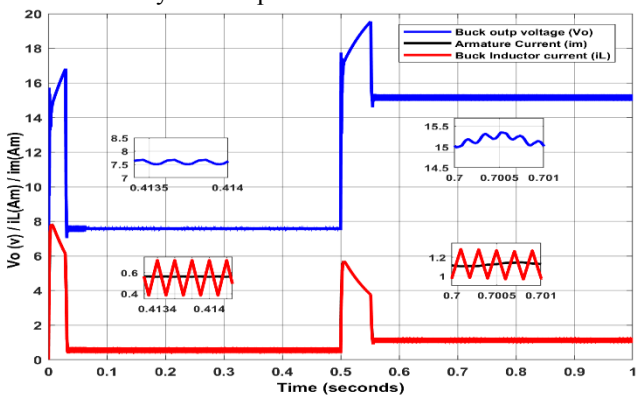


Fig 6. The outputs voltage vo, inductor current i_L, and armature current i_m

At the value of $T_L = 40m\ N.m$ it is seen in the Fig 6. that the vo is about $7.7v\ to\ 15.3v$, i_L and i_m are approximately between $0.56A$ to $1.1A$.

VI. CONCLUSION

This study was presented as a constrained MPC control for a DC-DC buck converter driven DC-motor, in which the discrete averaged model was used over the prediction horizon for the MPC control design. The quadratic programming (QP) optimization algorithm subjected to inequal constraints solved in every sampling interval to predict the right input control signal which used to govern the buck converter depending on the desired speed. The results illustrated that the DC-DC buck converter driven DC-motor could be handled successfully by the MPC controller. Also, the system output signal response was fast enough to reach the desired tracking trajectory. Future work could apply this MPC strategy to a nonlinear DC-DC buck-boost converter driven DC-motor.

REFERENCES

[1] L.Wang, S. Chai, D. Yoo, L. Gan, K. Ng, "PID and Predictive Control of Electrical Drives and Power Converters Using MATLAB/Simulink", published 2015 john wiley & sons singaphore pte. Ltd, (ISBN: 9781118339442)

[2] M.A. Ahmad, R.M.T. Raja Ismail, M.S. Ramli," Control Strategy of Buck Converter Driven Dc Motor: a Comparative Assessment", Australian Journal of Basic and Applied Sciences, 4(10): ISS 1991-8178

[3] J. Yang, H. Wu, L. Hu, S.Li, "Robust Predictive Speed Regulation of Converter-Driven DC Motors via a Discrete-Time Reduced-Order GPIO", IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS,VOL.66,NO.10, OCT 2019

[4] J.Linares-Flores, J. Reger, and H. Sira-Ram´ırez "Speed-sensorless tracking control of a DC-motor via a double Buck-converter", Proceedings of the 45th IEEE Conference on Decision & Control Manchester Grand Hyatt Hotel San Diego, CA, USA, December 13-15, 2006

[5] T. K. Roy, M. F. Pervej, F. K. Tumpa, and L. C. Paul, "Nonlinear Adaptive Controller Design for Velocity Control of a DC Motor Driven by a DC-DC Buck Converter Using Backstepping Ap", 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)8-10 December 2016, Rajshahi-6204, Bangladesh

[6] Ramon Silva-Ortigoza, Victor Manuel Hern ´ andez-Guzm ´ an, Mayra Antonio-Cruz, and Daniel Mu ´ noz-Carrillo, "DC/DC Buck Power Converter as a Smooth Starter for a DC Motor Based on a Hierarchical Control", IEEE TRANSACTIONS ON POWER ELECTRONICS, VOL. 30, NO. 2, FEBRUARY 2015

[7] Paul Krause, Oleg Wasynczuk, Scott D. Sudhoff, Steven Pekarek, "ANALYSIS OF ELECTRIC MACHINERY AND DRIVE SYSTEMS", IEEE Press 445 Hoes Lane Piscataway, NJ 08854 IEEE Press Editorial Board 2013 John Anderson, Editor in Chief

[8] M.A. Ahmad, R.M.T. Raja Ismail, M.S. Ramli," Speed Control of Buck-converter Driven Dc Motor Using LQR and PI: A Comparative Assessment", 2009 International Conference on Information Management and Engineering

[9] E. Lucas N, D. Helder R, D. Grace S, G. Jose C, "DC motor speed control via buck-boost converter through a state dependent limited frequency switching rule", 2017 IEEE 56th Annual Conference on Decision and Control (CDC), 2017-12, p.2072-2077

[10] Grace S. Deaecto, Jose C. Geromel, "Stability Analysis and Control Design of Discrete-Time Switched Affine Systems", IEEE TRANS ON AUTOMATIC CONTROL, VOL. 62, NO. 8, AUG 2017

[11] F. E. H. Velasco, J. E. Candelo-Becerra, A. R. Santamaría, "Dynamic Analysis of a Permanent Magnet DC Motor Using a Buck Converter Controlled by ZAD-FPIC"

[12] E. H. Márquez, R. S. Ortigoza, J. R. G. Sánchez,V. H. G. Rodríguez and J. N. A. Juárez, "A New DC/DC Buck-Boost Converter–DC Motor" System: Modeling and Experimental Validation", IEEE LATIN MERICA TRANSACTIONS, VOL. 15, NO. 11, NOV. 2017 2043

[13] L. Cheng, P. Acuna, Ricardo P. Aguilera, J. Jiang, S. Wei, John E. Fletcher, Dylan D. C. Lu "Model Predictive Control for DC–DC Boost Converters with Reduced-Prediction Horizon and Constant Switching Frequency", IEEE TRANSACTIONS ON POWER ELECTRONICS, VOL. 33, NO. 10, OCTOBER 2018

[14] Rodney H.G. Tan and Landon Y. H. Hoo, "DC-DC converter modeling and simulation using state space approach", 2015 IEEE Conference on Energy Conversion (CENCON), Electronic ISBN: 978-1-4799-8598-2, USB ISBN: 978-1-4799-8597-5

[15] Liuping Wang, "Model Predictive Control System Design and Implementation Using MATLAB", School of Electrical and Computer Engineering, RMIT University Melbourne, VIC 3000 Australia ISBN 978-1-84882-330-3 e-ISBN 978-1-84882-331-0

[16] B. Kouvaritakis and M. Cannon, "Model Predictive Control: Classical, Robust and Stochastic", Springer, 2015, (ISBN-13: 978-3319248516), (ISBN-10: 3319248510)

PCA, Random-Forest and Pearson Correlation for Dimensionality Reduction in IoT IDS

Alaa Alhowaide*
Computer Science Department
Memorial University of Newfoundland
St. John's, NL, Canada
azalhowaide@mun.ca

Izzat Alsmadi
Department of Computing and Cyber Security
Texas A&M University-San Antonio
San Antonio, Texas, United States
ialsmadi@tamusa.edu

Jian Tang
Computer Science Department
Memorial University of Newfoundland
St. John's, NL, Canada
jian@mun.ca

Abstract*— The exposure of Internet of Things (IoT) networks led to an exponential increase in traffic size. Thus, the traffic data meet the 5V big data model, which results in the rise of many cyber threats and the appearance of new security challenges. Intrusion Detection Systems (IDS) need to apply dimensionality reduction to handle the enormous data size and to speed-up cyber threats detection. This research analyzes the impact of Principal Component Analysis (PCA), Random-Forest (RF), and filter-based dimensionality reduction methods on several detection models in four different datasets. Additionally, it studies the impact of some parameters on dimensionality reduction methods' performance. Results showed that dimensionality reduction methods were able to reduce datasets' size without compromising the detection models' efficiency. Moreover, results showed that PCA was the best method in reducing datasets' size over all the datasets. At the same time, RF showed a better distinguishing ability with the minimum amount of information.

Keywords*— IoT, Feature Selection, Ensembling, Cyber-Security, Intrusion Detection.

I. Introduction

IoT networks are growing faster than expected. According to IoT Analytics [1], the number of connected devices reached 9.5 billion devices at the end of 2019. IoT Analytics expects this figure to reach 22 billion by 2025 [2]. This rapid growth came aligned with weak security measures and defenses, which has resulted in a significant increase in cyberattacks. Since the appearance of COVID-19, 71% of security professionals have noticed an increase in security threats or attacks according to research performed by Checkpoint [3]. Most of the attacks targeting IoT networks try to delay the service provided by the devices or to compromise and convert the devices into members of the botnet under the attacker's control to amplify the speed of their attacks.

An IDS, which refers to a device or software strategically allocated at a tactical point on a network to monitor all traffic passing through it [4], represents the first defense line against cyber-attacks. Machine Learning plays a crucial role in building IDSs to detect such threats. Network data is considered as big data as it meets the 5V big data features: volume, velocity, variety, veracity, and value. As a result, dimensionality reduction and feature selection steps are essential to overcome the poor performance of ML models' and explainability issues. Feature selection is a crucial step in the pre-processing data phase [5], defined as the process of selecting the minimal subset of features based on reasonable criteria with the aim to improve an algorithm performance [6]. Also, it is defined as the process of removing irrelevant, redundant, and noisy features [7]. Sometimes, it is referred to as the relevant feature identification process [8]. Feature selection provides an immediate impact on IDS performance in terms of computational complexity and accuracy. It selects only the most relevant and most distinguishing features between traffic types. Thus, it reduces the number of features to be considered by an IDS, and in turn, reduces the complexity of the problem [7]. Moreover, by eliminating the noisy, redundant, and useless features, the accuracy of IDS increases, and the generated number of false alerts decreases as well [8].

Dimensionality reduction refers to any method that reduces the number of input variables in a dataset. It includes feature selection, linear algebra methods, projection methods, and autoencoders [9]. Feature selection is classified into three main types: wrapper, filter, and hybrid methods. The wrapper methods try to optimize some predefined criteria concerning the feature set as part of the selection process. The filter methods rely on the general characteristics of the training data to select features that are independent of each other and are highly dependent on the class label. The hybrid methods try to exploit the features of both wrapper and filter methods [10].

This research aims to analyze the impact of dimensionality on the efficiency of detection models, and find the best dimensionality reduction method that reduces a dataset to its minimum with the lowest computational requirements. Furthermore, this research seeks to fill the research gap in the literature by not considering dimensionality dimensionality reduction methods for IoT IDS.

Section two provides a literature review. Section three explains the experiments, while section four discusses results. Lastly, section five concludes the study.

II. Literature Review

Many studies in literature have applied ML methods for detecting general botnet attacks. However, most of them did not use any kind of feature selection nor dimensionality reduction, such as the works [11], [12], [13], [14], [15]. On the other hand, some works applied only one feature selection method, such as [7] who used the Correlation Coefficient (CC), which measures

*Corresponding author.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

the correlation between features, as a filter to select features. CC is the simplest feature selection method to be used, where it aims to eliminate the most correlated features. When a correlation value close to +1 or -1, then these features are highly correlated and, thus, distracted from the final feature set. A zero-correlation value indicates no correlation between the features. The lowest N ranked features were selected as the most important features. [16] suggested two methods to select important features. The first method used the "leave-one-out" method with Naïve Bayesian classifier. The second method used the Gain Ratio technique. However, applying only one feature selection method raises doubts about how much the selected feature set represents the optimal feature set or at least how close we are to it. Few other studies have considered ensemble feature selection methods. [17] used Deep Feed Forward Neural Networks (DFFNN) to reduce the data dimensionality by generating embedded features. The embedded features were the inner last layer after removing the output layer from the DFFNN. In [18], a two-step hybrid approach was proposed. The first step used Pearson and Fisher filter to select a feature subset. The second step used the output of the first step as input for a wrapper method. The wrapper step used the K-Nearest Neighbors (K-NN) and Random Forest Tree (RFT) to select features based on the F1 score. Furthermore, the wrapper step considered the Sequential Forward Feature Selection (SFFS) and Sequential Backward Feature Elimination (SBFE) with both classifiers. In [19] a linear combination of multiple correlation filter measures were used. Examples of those measures include: correlation-feature-selection (CFS), minimal-redundancy-maximal-relevance ($mRMR$), Mahalanobis distance, and others. [20] proposed a dimensionality reduction method called features pairing, which is based on combining two feature selection methods. Feature pairing stands for generating a new feature represented as a string by concatenating the paired features' strings. The first pairing method was based on a wrapper that uses F-Measure (FM), and the second method was based on filters such as Information Gain Ratio (IGR). This method is computationally expensive, where it suggests computing the FM and IGR for all possible pairings. Furthermore, the need to extract the original features during runtime is an essential task, which means that the time reduction is not optimum. Few studies have considered feature selection matters in IDS, and only a small range of feature selection methods have been used, most of them based on correlation measures.

III. Experiments

Experiments were conducted on Compute Canada at Cedar cluster. All experiments used Broadwell type nodes, which are equipped with Intel(R) Xeon(R) CPU E5-2683, v4, 2.10GHz, and 125G RAM. Sklearn 0.23.2, and Python 3.7.4 were used to implement all the experiments. tThis research considered the detection problem as a binary classification problem to develop an IDS for general-purpose detection. In addition, the experiments considered two performance measures are F-Score and ROC-AUC score. These metrics provide an informative view of the effectiveness of the detection model. These metrics are calculated as below:

The F-Score represents the trade-off between the precision and recall regarding the positive class by using the harmonic mean. The best model has a value of 1.

$$F - Score = 2 * \frac{Recall\ X\ Precision}{Recall + Precision} = \frac{2 * TP}{(2 * TP + FN + FP)}$$

ROC-AUC represents the Area Under the Receiver Operating Characteristic Curve. By computing the area under the ROC curve, the curve information is summarized in one number. When the integral boundaries are reversed as large T has a lower value on the x-axis, then the AUC is given by:

$TPR(T): T \rightarrow y(x)$

$FPR(T): T \rightarrow x$

$$A = \int_{x=0}^{1} TPR(FPR^{-1}(x))dx$$

Moreover, all the results presented used the mean of the five folds of the cross-validation.

A. Datasets

This research utilized four datasets NSL [21], NB15 [22], BoTNetIoT [11], and BoTIoT [23] datasets. NSL dataset is the most commonly used dataset in intrusion detection, which contains regular network traffic. Similarly, NB15 contains regular network traffic. NSL and NB15 represent a benchmark for intrusion detection. They also enable the current research to be compared with previous works. BoTNetIoT dataset contains the traffic of nine IoT devices. Mirai and Gafgyt botnet attacks are the only attacks in the BoTNeTIoT dataset. BoTNeTIoT dataset consists of 23 extracted features over five different time-windows. In [24], it was indicated that the best classification performance was at time-window 10 seconds (0.1 decay factor). Thus, our research considered the 23 features at a 10 second time-window only. BoTIoT is a recently published dataset of a simulated IoT network traffic. BoTIoT has a variety of recent/new attacks. This research concedered BoTNeTIoT and BoTIoT datasets because they include veraiety of recent attacks. This research considered BoTNeTIoT and BoTIoT datasets because they are the only IoT datasets that include a variety of recent IoT attacks.

B. Feature selection methods

Different dimensionality dimensionality reduction methods use different ways to reduce dataset size. This research considered the three dimensionality dimensionality reduction methods: PCA, RF, and Top-*n*, where *n* represents the number of features to select. Below is an explanation of each method and how it works:

- PCA: It is a linear dimensionality reduction method that uses Singular Value Decomposition of the data to project it to a lower-dimensional space. PCA compresses all features into a smaller number of dimensions. Thus, it still requires an IDS to extract all features, which will result in spending additional time in the detection process [25]. Moreover, as the number of dimensions generated decreases, more descriptive

TABLE 1 Dimensionality reduction methods datasets size reduction, and number of features selected.

	NSL		NB15		BoTIoT		BoTNeTIoT	
Dimensionality Reduction Method	**Size Reduction (%)**	**# Features**	**Size Reduction (%)**	**# Features**	**Size Reduction (%)**	**# Features**	**Size Reduction (%)**	**# Features**
All	0	41	0	43	0	43	0	23
PCA-0.95	0.9999	8	0.9999	12	0.9999	8	0.9999	5
PCA-0.97	0.9999	10	0.9999	14	0.9999	9	0.9999	7
PCA-0.99	0.9999	15	0.9999	19	0.9999	11	0.9999	9
RF-1	0.8749	4	0.8372	6	0.8139	6	0.7916	3
RF-10	0.7999	7	0.6976	12	0.8372	5	0.6666	6
RF-100	0.6999	11	0.6511	14	0.8372	5	0.7499	4
Top-5	0.8499	5	0.8604	5	0.8604	5	0.7499	5
Top-10	0.7249	10	0.7441	10	0.7441	10	0.5416	10

information will be lost. Choosing the number of components is significant, because it affects the reduction percentage and the method efficiency. Choosing a fixed number of components is not a proper or practical solution. Thus, we considered a practical way that uses the variance explained by each component. The variance stands for the amount of information represented by each set of components. As the number of components decreases the loss of information increases. Fig. *1* shows the variance percentage for different numbers of components over the NSL dataset. The variance percentage increases as the number of components increases. Using 40 components for the NSL dataset means that no size reduction is acquired, where each component represents a feature. Typically, we want the explained variance to be between 95–99%. Thus, we considered three different variance percentages are 95, 97, and 99%.

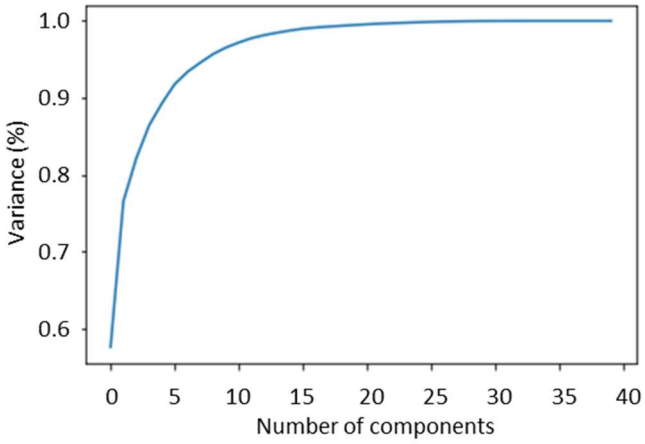


Fig. 1. PCA explained variance vs the number of components.

- RF: It is an ensemble feature selection method. It ensembles the number of decision tree classifiers on various sub-samples of the dataset and works as a meta-estimator. It then uses averaging to improve the predictive accuracy and control over-fitting. It uses bootstrapping for sub-sampling. This research used SciKit-Learn [26] implementation that uses the square-root of the total number of features as the maximum number of features to extract [27]. After getting the maximum number of features, only those features with support scores higher than the ID feature's support score were selected. Moreover, as the number of trees within the RF has a considerable impact on its performance, thus we considered constructing forests of one, 10, and 100 trees.
- Top-*n*: It is a univariate feature selection that works by selecting the best features based on univariate statistical tests. It is also known as a filter-based feature selection method. This method uses the Pearson correlation measure to rank the features. Then, it selects the top *n* features.

IV. Results

TABLE *1* shows the obtained percentage of size reduction and the number of chosen features for each dataset. The "All" method stands for instances when all the features were used without any reduction. PCA-*X* stands for instances when the PCA method was applied with a total variance percentage equals to *X*. For example, PCA-0.95 stands for using PCA with an accumulative variance percentage equals to 0.95. Similarly, RF-*X* and Top-*n* stand for using RF composed of *X* trees, and top *n* features, respectively. PCA had a maximum reduction percentage over all the datasets. PCA reduction percentage was the same using different variance percentages. Noticeably, RF reduction percentage decreased as the number of inner trees increased over all the datasets.

In Fig. *2* to Fig. *5*, the F-scores over each dataset is displayed. The ROC-AUC score over each dataset is displayed in Fig. *6* to Fig. *9*. From Fig. *2*, it can be seen that the highest F-score is generated by KNN and RF classifiers using all features, all the PCAs, and RF-100. The distinguishing ability of the PCAs methods is stronger than the RFs methods in the NSL dataset. Additionally, all PCAs methods achieve scores equal to or higher than using all features. The Top-*n* methods show the

lowest scores. The Top-10 shows the same scores as Top-5 for all classifiers, except for GaussianNB. Nearly all PCAs have the same scores overall variance percentages. Additionally, when using BernoulliNB the RFs scores increase as the number of trees increases. The reverse is true when using GaussianNB and DT.

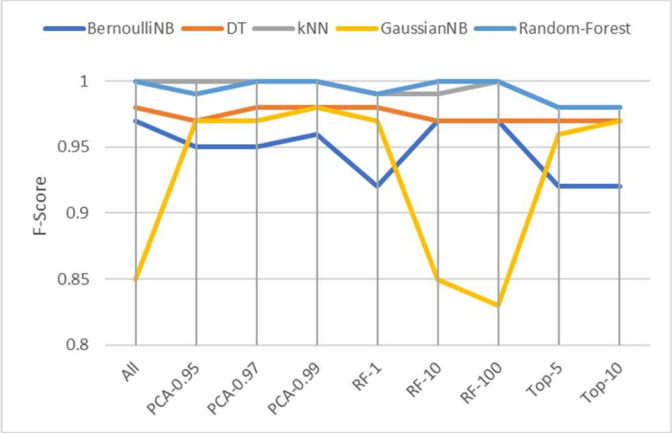


Fig. 2. NSL F-Score.

In Fig. *3*, which displays the F-scores using the NB15 dataset, the RFs methods show higher scores than using all feature, PCA-0.95, and PCA-0.97 for overall classifiers, except GaussianNB classifier. The PCA-0.99 show the highest scores over BernoulliNB, KNN, and GaussianNB. The Top-*n* methods show the same performance using DT, KNN, and RF. The Top-5 overcome Top-10 using GaussianNB, and vice versa when using BernoulliNB. The PCAs scores increase as the variance increases using BernoulliNB, and vice versa using GaussianNB. Notably, the number of trees within an RF does not show any impact using all classifiers, except using DT and Gaussian, the scores decrease as the number of trees increases.

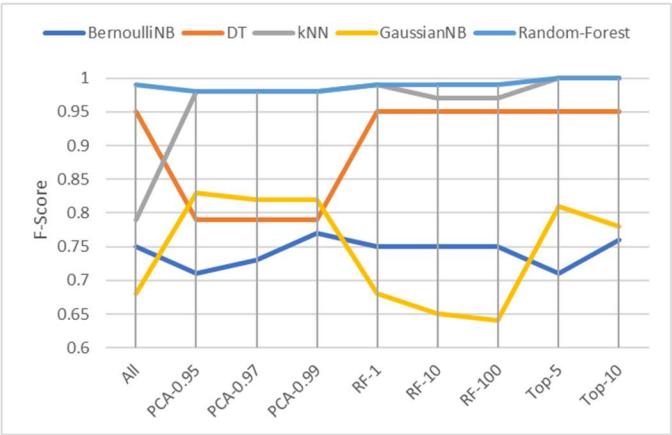


Fig. 3. NB15 F-Score.

Fig. *4* shows the F-scores using the BoTIoT dataset. The BernoulliNB is removed from the figure because it generated zero scores using the overall methods. Significantly, the RFs methods out-performed PCAs using all the classifiers. The RFs shows score equal and bellow the scores when using all the features. The Top-5 out-performs Top-10 using DT and RF, and vise versa when using KNN and GaussianNB. Notably, PCAs' scores increase as variance increases using DT, kNN, and RF. On the other hand, they decrease when using gaussianNB. Generally, RFs' scores increase as the number of trees increases.

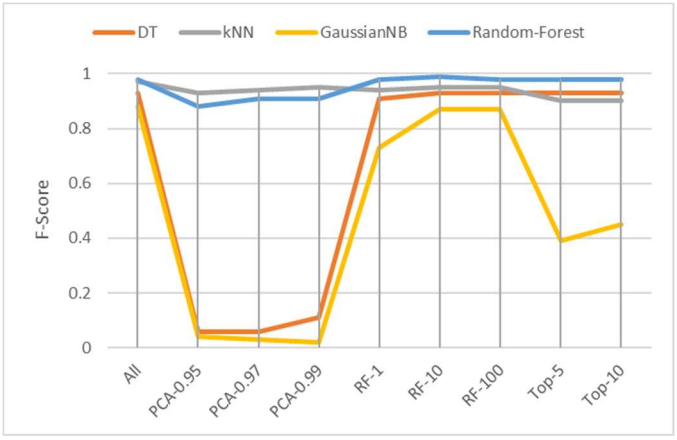


Fig. 4. BoTIoT F-Score.

The BoTNeTIoT dataset F-scores are displayed in Fig. *5*. The GaussianNB shows a zero score using all features. Significantly, the PCAs out-perform using all features using BernoulliNB. Moreover, the PCAs show the same scores as RFS, except for BernoulliNB, the PCAs overcome the RFs methods. The Top-*n* methods show a zero score when using BernoulliNB. Additionally, they show the same scores as PCAs for all classifiers. Notably, the PCA scores increase as variance increases. Also, the number of trees does not have an impact on RFs' scores.

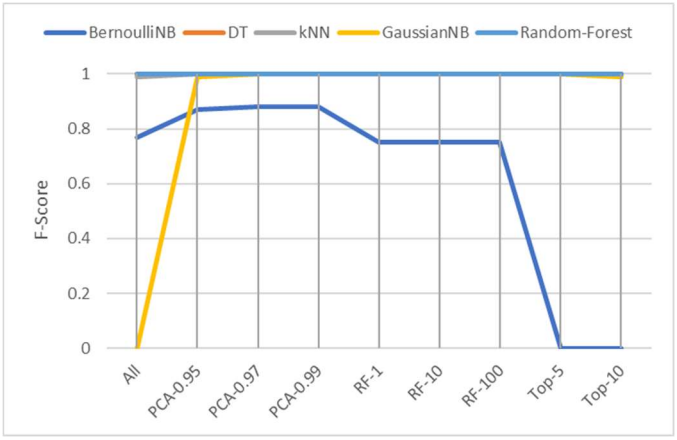


Fig. 5. BoTNeTIoT F-Score.

In Fig. *6*, the NSL ROC-AUC scores are displayed. The RFs scores are equal to or higher than the PCAs' scores. Similarly, the RFs scores and using all features. Notably, Top-10 overcome Top-5 score using BernoulliNB. The PCAs and RFs scores slightly affected by variance percentage and number of trees, respectively.

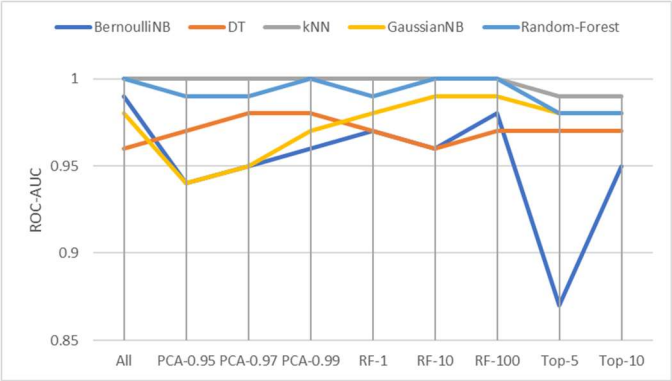


Fig. 6. NSL ROC-AUC score.

In Fig. 7, PCAs overcome RFs using BernoulliNB and GausianNB, vice versa using DT and RF. The PCAs overcome using all features using BernoulliNB and GaussianNB. Additionally, all methods overcome using all features using kNN. The Top-5 overcome Top-10 for all classifiers, except for BernoulliNB. Generally, there is no impact of variance on PCAs' scores. The RFs' scores increase as the number of trees increases using BernoulliNB, and vice versa using GaussianNB.

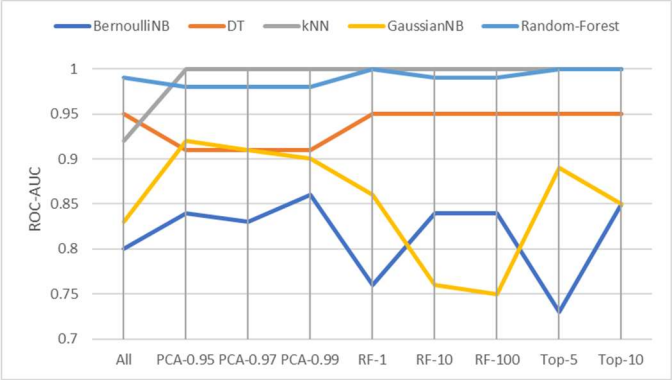


Fig. 7. NB15 ROC-AUC score.

Fig. 8 All methods have the same score using kNN and GaussianNB. Similarly, All methods, except PCAs and RF-10, have the same score using DT. Additionally, the PCAs overcome using all features using BernoulliNB. Also, the RFs slightly overcome using all features using RF classifier. As can be seen, the PCAs' scores increase as the variance increases using RF classifier. The number of trees does show an impact on the RFs' scores.

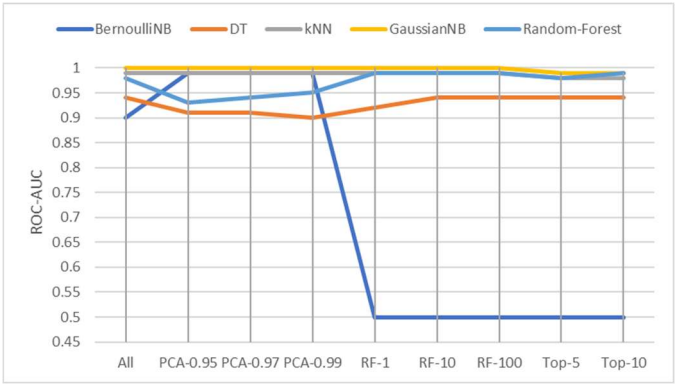


Fig. 8. BoTIoT ROC-AUC score.

All methods have the same scores except for BernoulliNB in Fig. 9. Significantly, the PCAs and RF-10 show the highest scores and overcome using all features using BernoulliNB. Furthermore, The Top-10 overcomes Top-5 using BernoulliNB. The variance nor the number of trees has an impact on the PCAs or RFs, respectively.

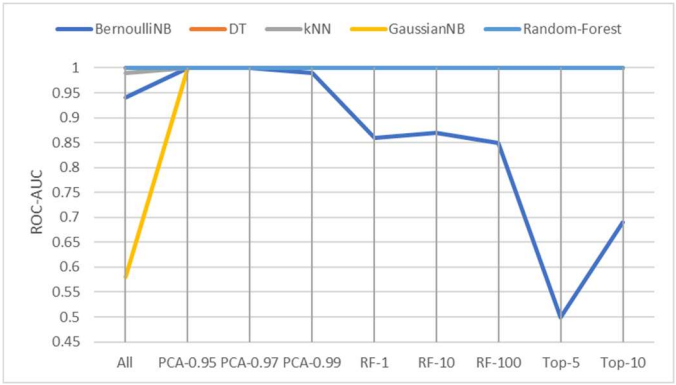


Fig. 9. BoTNeTIoT ROC-AUC score.

V. Conclusion

Dimensionality reduction reduces the noise and redundancy in data. As a result, classifiers' efficiency is improved, as the experiments show. In some cases, the data itself has the most significant impact on classifiers' efficiency more so than any other factor.

Classifiers that depend on trees in classification show a better performance using RF methods. This impact can be a sign of overfitting of the model on the selected features. Similarly, when considering the PCA with BernoulliNB and GaussianNB classifiers.

Generally, the impact of the variance percentage of PCA and the number of RF trees do not have a significant impact on classifiers' efficiency. Thus, using a low variance percentage generates lesser dimensions, which save more time during detection/classification without significantly compromising model efficiency—similarly, the number of trees has the same impact in RF.

PCA shows the highest size reduction, but it still requires to extract all features from the traffic. Thus, the PCA consumes additional time over using RF.

REFERENCES

[1] "IoT 2019 in Review: The 10 Most Relevant IoT Developments of the Year." https://iot-analytics.com/iot-2019-in-review/ (accessed May 27, 2020).

[2] "State of the IoT 2018: Number of IoT devices now at 7B – Market accelerating." https://iot-analytics.com/state-of-the-iot-update-q1-q2-2018-number-of-iot-devices-now-7b/ (accessed May 27, 2020).

[3] "A Perfect Storm: the Security Challenges of Coronavirus Threats and Mass Remote Working," *Check Point Software*, Apr. 07, 2020. https://blog.checkpoint.com/2020/04/07/a-perfect-storm-the-security-challenges-of-coronavirus-threats-and-mass-remote-working/ (accessed May 27, 2020).

[4] T. Mohamed, T. Otsuka, and T. Ito, "Towards Machine Learning Based IoT Intrusion Detection Service," *Recent Trends and Future Technology in Applied Intelligence. IEA/AIE 2018. Lecture Notes in Computer Science*, vol. 10868, May 2018, doi: https://doi.org/10.1007/978-3-319-92058-0_56.

[5] A. Blum and P. Langley, "Selection of Relevant Features and Examples in Machine Learning," *Artificial Intelligence,* vol. 97, pp. 245–271, 1997.

[6] H. Liu and H. Motoda, *Feature Selection for Knowledge Discovery and Data Mining*, vol. 454. Springer Science & Business Media, 2012.

[7] N. Moustafa, B. Turnbull, and K.-K. R. Choo, "An Ensemble Intrusion Detection Technique Based on Proposed Statistical Flow Features for Protecting Network Traffic of Internet of Things," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4815–4830, Jun. 2019, doi: 10.1109/JIOT.2018.2871719.

[8] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network Anomaly Detection: Methods, Systems and Tools," *IEEE Communications Surveys Tutorials*, vol. 16, no. 1, pp. 303–336, First 2014, doi: 10.1109/SURV.2013.052213.00046.

[9] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Elsevier, 2011.

[10] Y. Chen, Y. Li, X.-Q. Cheng, and L. Guo, "Survey and Taxonomy of Feature Selection Algorithms in Intrusion Detection System," *International Conference on Information Security and Cryptology, Springer*, pp. 153–167, 2006.

[11] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection," *arXiv:1802.09089 [cs]*, May 2018, Accessed: Oct. 24, 2019. [Online]. Available: http://arxiv.org/abs/1802.09089.

[12] M. AL-Hawawreh, N. Moustafa, and E. Sitnikova, "Identification of malicious activities in industrial internet of things based on deep learning models," *Journal of Information Security and Applications*, vol. 41, pp. 1–11, Aug. 2018, doi: https://doi.org///doi.org/10.1016/j.jisa.2018.05.002 ID: 287016.

[13] A. Verma and V. Ranga, "ELNIDS: Ensemble Learning based Network Intrusion Detection System for RPL based Internet of Things," in *2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU)*, Apr. 2019, pp. 1–6, doi: 10.1109/IoT-SIU.2019.8777504.

[14] M. A. Jabbar, R. Aluvalu, and S. S. S. Reddy, "Cluster Based Ensemble Classification for Intrusion Detection System," in *Proceedings of the 9th International Conference on Machine Learning and Computing - ICMLC 2017*, Singapore, Singapore, 2017, pp. 253–257, doi: 10.1145/3055635.3056595.

[15] S. T. Miller and C. Busby-Earle, "Multi-Perspective Machine Learning a Classifier Ensemble Method for Intrusion Detection," in *Proceedings of the 2017 International Conference on Machine Learning and Soft Computing - ICMLSC '17*, Ho Chi Minh City, Vietnam, 2017, pp. 7–12, doi: 10.1145/3036290.3036303.

[16] N. T. Pham, E. Foo, S. Suriadi, H. Jeffrey, and H. F. M. Lahza, "Improving performance of intrusion detection system using ensemble methods and feature selection," in *Proceedings of the Australasian Computer Science Week Multiconference on - ACSW '18*, Brisband, Queensland, Australia, 2018, pp. 1–6, doi: 10.1145/3167918.3167951.

[17] Y. Zhou, M. Han, L. Liu, J. S. He, and Y. Wang, "Deep learning approach for cyberattack detection," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Apr. 2018, pp. 262–267, doi: 10.1109/INFCOMW.2018.8407032.

[18] A. Guerra-Manzanares, H. Bahsi, and S. Nõmm, "Hybrid Feature Selection Models for Machine Learning Based Botnet Detection in IoT Networks," in *2019 International Conference on Cyberworlds (CW)*, Oct. 2019, pp. 324–327, doi: 10.1109/CW.2019.00059.

[19] H. T. Nguyen, K. Franke, and S. Petrović, "A new ensemble-feature-selection framework for intrusion detection," in *2011 11th International Conference on Intelligent Systems Design and Applications*, Nov. 2011, pp. 213–218, doi: 10.1109/ISDA.2011.6121657.

[20] M. Milliken, Y. Bi, L. Galway, and G. Hawe, "Ensemble learning utilising feature pairings for intrusion detection," in *2015 World Congress on Internet Security (WorldCIS)*, Oct. 2015, pp. 24–31, doi: 10.1109/WorldCIS.2015.7359407.

[21] "NSL-KDD | Datasets | Research | Canadian Institute for Cybersecurity | UNB." https://www.unb.ca/cic/datasets/nsl.html (accessed Nov. 20, 2019).

[22] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *2015 Military Communications and Information Systems Conference (MilCIS)*, Nov. 2015, pp. 1–6, doi: 10.1109/MilCIS.2015.7348942.

[23] "The BoT-IoT Dataset." https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/bot_iot.php (accessed Dec. 12, 2019).

[24] A. Alhowaide, I. Alsmadi, and J. Tang, "Features Quality Impact on Cyber Physical Security Systems," in *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, BC, Canada, Oct. 2019, pp. 0332–0339, doi: 10.1109/IEMCON.2019.8936280.

[25] "sklearn.decomposition.PCA — scikit-learn 0.23.1 documentation." https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html (accessed Jul. 08, 2020).

[26] "scikit-learn: machine learning in Python — scikit-learn 0.23.1 documentation." https://scikit-learn.org/stable/index.html (accessed Jul. 08, 2020).

[27] "3.2.4.3.1. sklearn.ensemble.RandomForestClassifier — scikit-learn 0.23.1 documentation." https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html?highlight=random%20forest#sklearn.ensemble.RandomForestClassifier (accessed Jul. 08, 2020).

Baseline Assessment, Intervention, and Interval Reassessment of Clinical Information Systems Infrastructure in an Academic Regional Referral Emergency Department

Amelia L. Gurley
Alpert Medical School of Brown University
Providence, RI, USA
amelia_gurley@brown.edu

James E. Tanch MD
Alpert Medical School of Brown University
Providence, RI, USA
james_tanch@brown.edu

David C. Portelli MD
Alpert Medical School of Brown University
Providence, RI, USA
dportelli@lifespan.org

Cedric Priebe III MD
Lifespan
Providence, RI, USA
cedric.priebe@lifespan.org

Jeffrey A. Berube
Lifespan
Providence, RI, USA
jberube@lifespan.org

Adam Dieffenbach
Lifespan
Providence, RI, USA
adam.dieffenbach@lifespan.org

Renee Harris
Lifespan
Providence, RI, USA
renee.harris@lifespan.org

Maliha Jalal
Lifespan
Providence, RI, USA
mjalal@lifespan.org

Stephen L. Sayles
Lifespan
Providence, RI, USA
ssayles1@lifespan.org

Allan J. Smulling
Lifespan
Providence, RI, USA
asmulling@lifespan.org

Gregory D. Jay MD PhD
Alpert Medical School of Brown University
Providence, RI, USA
gjay@lifespan.org

Leo Kobayashi MD
Alpert Medical School of Brown University
Providence, RI, USA
lkobayashi@lifespan.org

***Abstract*—The widespread use of clinical information systems (CIS) creates ample opportunities for CIS-associated operational disruptions and delays that can negatively impact patient care. A clinical quality improvement (QI) project studied the CIS infrastructure at an academic regional referral emergency department (ED) at baseline and after implementation of targeted sustainment interventions. An ED 5S workgroup collaborated with ED clinical leaders and institutional information services (I.S.) teams to conduct comprehensive CIS inventory and function testing in 2016; an end-user survey on CIS functionality and impact were administered. Findings were analyzed to determine common points of failure, and concerted interventions were implemented during 2016-2018: development of a structured ED CIS inventory document with device mapping and specification of clinical functions; prioritized repair / replacement of non-functional / missing devices; institution of scheduled ED "Tech Rounds"; and installation of a self-service work-ticket Web portal for repair requests. A follow-up reassessment and survey in 2018 revealed mixed intervention results and highlighted infrastructure elements superseded by advances in mobile technology. Overall, a multi-disciplinary effort assessed an ED CIS infrastructure and its functionality at baseline and over time for changes associated with targeted interventions. Studied metrics revealed some successes and ongoing challenges.**

Keywords*— *emergency medicine, hospital information systems, information technology, medical informatics, quality improvement

I. Introduction

A. Background

Clinical information systems (CIS) are pervasive and indispensable to medical practice and health care delivery in the United States [1]. Encompassing health information technology (HIT) devices, programs, interfaces, and processes which collect, manipulate, store, retrieve, and transfer information for clinical purposes at the point of care, CIS in effect became inextricably embedded into the U.S. health care system with the passage of the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009 [2]. This widespread adoption of HIT and CIS has demonstrably introduced new complexities into patient care at the bedside [3-17]. Not unexpectedly, reports of system errors, inter-operability challenges, network outages, hazardous user interfaces, and other CIS issues abound in the literature [18-23]; reports of patient harm exist [24-25].

B. Importance

Emergency Departments (EDs) operate uninterruptedly on a 24/7/365 basis to evaluate and manage any and all presenting patients– this makes EDs some of the most continually information-intensive spaces in a medical system. Combined with the frequently time-critical, severe, and complex nature of ED patient illnesses and injuries, any delays, interruptions, and errors introduced by CIS can become particularly dangerous. CIS has been listed as one of the key targets for improving ED front-end patient processing [26], and technology inadequacies have explicitly been cited as a source of dangerous ED medical errors [3, 27]. Slowdown time is cited by several studies as among the primary challenges that CIS use causes clinicians [28-32]; one 2005 study of provider workflow classified over 60% of all studied clinical interruptions as caused by "technology" [33]. The situation is compounded when on-duty clinical providers are relied on to recognize, diagnose, and troubleshoot CIS issues in real-time while performing primary clinical duties. Furthermore, ED facilities have frequently needed to deploy new CIS into pre-existing physical spaces and clinical infrastructure, an incremental process that is exacerbated by the fragmented nature of CIS products and the HIT landscape overall [34-35].

C. Goals of the Investigation

Multi-disciplinary efforts to document, address, and prevent

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

challenges associated with ED CIS infrastructure have not been widely reported. In order to better understand the functionality and problems associated with an ED CIS infrastructure, clinical providers in an ED 5S workgroup,[1] institutional information services (I.S.) personnel, and administrative leaders at an academic regional referral center initiated a collaborative quality improvement (QI) project. Specific project objectives were 1.) to generate a core ED CIS infrastructure knowledgebase with evaluative methods that employed objective metrics and 2.) to explore corrective and preemptive intervention efforts with before-and-after comparisons.

II. Methods

A. Project Setting

The project team conducted a preliminary CIS inventory and function testing of the project ED in February 2016. This inventory assessed the presence, location, and functionality of CIS devices managed by I.S. including stationary workstations, mobile workstations, printers, tap-in/out user authentication devices, digital signature capture devices, fingerprint scanners, and dictation microphones. Data center ("cloud-based") assets, biomedical engineering systems, and devices designated for non-clinical personnel were excluded. Testing was limited to functions that did not access patient records or health information.

The preliminary inspections noted several failure points relating to ED CIS infrastructure: 1.) the absence of a unifying and comprehensive knowledgebase of clinical workstation and device make / model / type, name, location, connectivity, userbase (in turn specifying required functions and functionality), known problems; 2.) operational barriers associated with reporting ED CIS issues; and 3.) the need for preemptive, recurring ED CIS maintenance procedures with dedicated human resources.

B. Project Interventions, Outcomes, and Analyses

After a collaborative review of the preliminary inventory data to establish a standardized checklist process for all ED CIS bedside devices and functions, a formal inventory and mapping assessment was conducted in April 2016 to set a baseline for longitudinal monitoring. Concerted QI interventions were then engineered by ED 5S and I.S. personnel for implementation to address issues deemed likely to be high-yield by the ED clinical leadership. Specific interventions included the following items and processes: 1.) creation and maintenance of a structured mapped inventory document with corresponding device labeling and assessments of specified essential functions, 2.) high-priority repair / replacement of malfunctioning / absent ED CIS devices, 3.) establishment of weekly scheduled ED "Tech Rounds" by a designated and I.S.-empowered administrative assistant, and 4.) institution-wide installation of an online HIT self-service Web portal (HEAT system; Ivanti, South Jordan, UT) to allow rapid, user-friendly submission of service requests. A complementary Web survey for end-user feedback was administered to query ED clinical personnel on their understanding and perceptions of ED CIS functionality and available technical support as well as the effect of ED CIS issues on patient care delivery at the project site.

The ED CIS infrastructure was re-assessed in April 2018 with repeat inventory and mapping, repeat Web survey, and analysis of the I.S. log record containing all submitted ED CIS work-ticket[2] data for the intervening period. Work-tickets were sorted into categories based on the main content of the request (including a category of "not CIS"); non-blinded thematic analyses were completed by a project member (ALG) using NVivo software (QSR International, Melbourne, Australia).

The primary outcome metric used to gauge the effect of project interventions was the measured differences between pre-intervention functionality and post-intervention functionality of categorical ED CIS elements, e.g., proportion of clinical work-stations with critical failures; proportion of mobile workstations with e-signature capture capability. Secondary outcome metrics included provider-reported changes in the following areas: perceived ED CIS device functionality; perceived ED CIS impact on clinical duties; the number and types of work-tickets submitted to I.S.; and clinical provider workflow surrounding work-ticket submission. ED CIS workstation replacement was monitored.

III. Results

A. ED CIS Infrastructure Inventory, Mapping, and Function Testing

1) Baseline Assessment

The project's preliminary ED CIS inventory in February 2016 examined 134 clinical provider workstations (out of 143 total ED workstations[3]) comprising 106 desktop workstations (77 PCs; 29 thin clients [TC]) and 27 mobile workstations-on-wheels (1 of 28 not able to be located). Critical performance issues were noted for 1% of fixed workstations (1 non-booting TC) and 15% of mobile workstations (3 non-booting devices; 1 device without network connectivity). Based on collaborative review of initial findings and characterization of failure points, a project checklist was derived for core workstation features (powering on, adequate compute / processing power, access to electronic medical record [EMR] system, up-to-date system security status) and functions (tap-in/-out logon, fingerprint-based authentication, e-signature capture [mobile workstations only], real-time dictation transcription, network and print connectivity).

The project team then completed a formal ED CIS inventory in April 2016 and generated a baseline, structured mapped inventory document– this compilation contained tabular information and graphical maps that integrated device specifications, name, location, network information, and primary userbase (with matched core features and functions), functionality, and identified problems (see Fig.1). Using the project checklist, 135 clinical provider workstations were fully inventoried and mapped– 2 (2%) of 106 fixed workstations and 7 (24%) of 29 mobile workstations were found to have critical performance issues. All devices tested were properly secured with institution-required security software. Specified ED CIS functions were variably functional at workstations requiring them, e.g., tap-in/-out modules were missing at 11% of 92 fixed workstations (with

[1] 5S refers to a specific workplace organization method based on "sort," "set in order," "shine," "standardize," and "sustain" [36].

[2] HIT support typically requires frontline end-users to initiate "pull" requests that generate work-tickets for repair, replacement, and maintenance services; remote assistance and mobile apps present methodologic improvements to the existing approach.

[3] Nine ED workstations were designated for use by support / ancillary staff, e.g., interpreters, and non-hospital personnel, e.g., EMS providers.

6% of 82 installed modules non-functional) and 32% of 25 mobile workstations (24% of 17 installed modules non-functional). Dictation microphones were absent from 1 (2%) of 44 dictation-enabled workstations. Printing from the EMR system failed at 5% of fixed workstations (see Table 1 for details).

2) Targeted Interventions

All inventoried devices were labeled with laminated stickers to facilitate their identification for work-ticket submission. Workstations with critical device issues were repaired or replaced with high priority by I.S. teams. Tap-in/-out modules, fingerprint readers, e-signature capture pads, and dictation microphones were repaired or installed as necessary; printpaths between workstations and printers were precisely specified and locked (where possible) for all ED provider workstations. ED "Tech Rounds" were conducted from April 2016 for several months until put on hold due to budgetary constraints (these resumed in 2019). The HEAT system went online in August 2016.

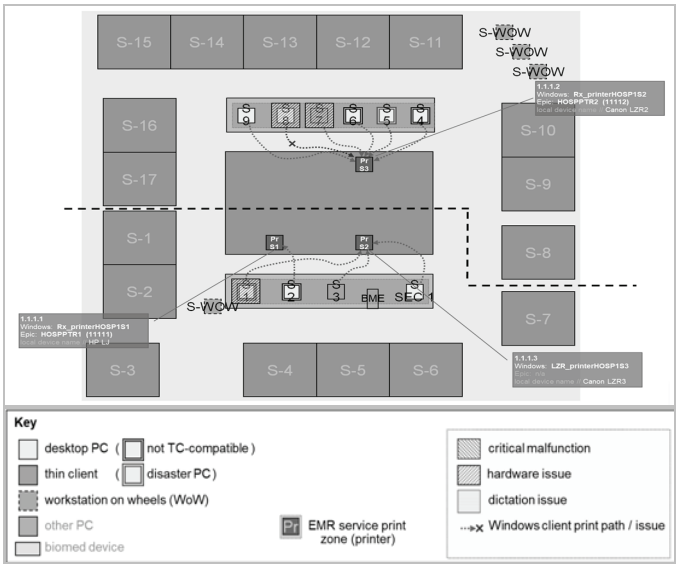


Figure 1: Example diagram of ED CIS device, function, and functionality mapping.

3) Follow-up Assessment

The project's 2018 reassessment examined 131 workstations, with 105 fixed workstations (76 PCs; 29 TCs) and 26 mobile workstations-on-wheels.[4] Critical performance issues were noted for 1 (1%) fixed workstation and 1 (4%) mobile workstation. All devices tested remained properly secured with institution-required security software. Tap-in-/out modules were missing at 31% of 101 fixed workstations (with 2% of 70 installed modules non-functional) and 38% of 25 mobile workstations (6% of 16 installed modules non-functional). E-signature capture was functional at 85% of 26 mobile workstations. Fingerprint authentication devices and dictation microphones were replaced in effect by smartphone-based mobile solutions. Printing from the EMR failed at 6% of fixed workstations.

Comparison of pre-intervention functionality and post-intervention functionality of categorical ED CIS elements revealed mixed intervention results and highlighted infrastructure elements superseded by advances in mobile technology (see Table 1 for details). From 2016 to 2018, a total of 43 fixed workstations and 15 mobile workstations were replaced, representing a replacement rate of 44% of devices over two years.

B. *End-user Web Surveys*

The 2016 Web survey was completed by 108 (22%) respondents out of 495 active advanced practice providers, nurses, physicians, scribes, and technicians; most respondents were nurses (39%) or physicians (40%). General CIS functions such as patient registration, chart access and order entry, laboratory and imaging results access and review, and Web access were reported to work well / consistently; dictation and printing were reported to work poorly / inconsistently. End-users indicated that a mean of 25 ± 27% of their shifts in the past month had been hindered by an ED CIS issue; comments attributed this to the absence of routine maintenance and to the inability to rapidly request necessary repairs and replacements. Service requests to I.S. were reported as initiated for a mean of 13 ± 2% ED CIS issues, predominantly through telephone and unit secretary delegation; 58 ± 38% of work-tickets were perceived to result in corrective action (see Table 2).

The follow-up Web survey in 2018 elicited 219 (44%) responses from 501 eligible ED personnel; most responses were from physicians (40%) or nurses (42%). This survey showed no significant change in perceived ED CIS functions except for dictation (p=0.005). Respondents reported that 29 ± 28% of shifts in the past month were significantly hindered by CIS issues (NS). The survey indicated an increase in I.S. work-ticket submissions (20 ± 28%; p=0.028), of which 2% were submitted by self-service Web portal (see Table 2).

C. *Work-ticket Data*

Thematic analysis of 2,062 I.S. work-tickets submitted between September 2014 and April 2018 identified 42% as consisting of non-CIS issues. Of the 1,187 work-tickets studied, the majority related to general system functions, printing, and inter-/intra-net access. Changes in ED CIS work-ticket types and time-to-work-ticket-completion could not be associated with project interventions (see Table 3 and Fig. 2).

IV. LIMITATIONS

The team's inability to access live EMR datastreams impaired project assessments. Changes at the institution level affected ED CIS independently from project interventions. Survey response rates doubled between baseline and repeat assessments and may have reflected different respondent samples. The longitudinal monitoring of CIS infrastructure between baseline and follow-up assessments was limited by use of work-ticket surrogate markers.

V. DISCUSSION

An ED 5S workgroup of frontline clinical providers worked with ED leaders and I.S. teams on a multi-disciplinary effort to assess the dynamic landscape of an Emergency Department Clinical Information System over time. Compiling a de novo knowledgebase with immediate operational utility and exploring processes to assist with the sustainment of critical infrastructure elements which support clinical patient care, the workgroup examined ED CIS from a sharp-end perspective. Several QI interventions were implemented to address distinct failure points; mixed results were obtained through project interventions and parallel developments.

[4] Due to ED re-configuration 2016-2018, some workstations were removed.

Table 1: Inventory of CIS device functionality.

ED Clinical Provider Workstation Inventory	February 2016 (Preliminary)	% upgraded relative to previous inventory*	April 2016 (Pre-intervention baseline)	% upgraded relative to previous inventory*	April 2018 (Post-intervention follow-up)	% upgraded relative to previous inventory*
Workstations in ED	143	-	147	-	145	-
Non-clinician workstations (*e.g.,* EMS workstation)	9	-	9	-	9	-
ED clinical provider workstations	134	-	138	-	136	-
Workstation not able to be located (*e.g.,* undergoing repair; pending installation)	1	-	3	-	5	-
ED clinical provider workstations inventoried	**133**	-	**135**	-	**131**	-
Fixed / desktop workstations	106		106		105	
-Standard desktop PC	72	n/a	72	1%	72	41%
-Thin client	29		29		29	
-Downtime desktop PC	5		5		4	
Mobile workstations	27	-	29	0%	26	58%

* Investigators were unable to determine the percentage of workstations replaced with an identical workstation model.

ED Clinical Provider Workstation Function Testing	February 2016 (Preliminary)	April 2016 (Pre-intervention baseline)	April 2018 (Post-intervention follow-up)	Interval Change (% difference between pre-intervention baseline and post-intervention follow-up states)	Comment
Fixed / desktop workstations	n=106	n=106	n=105		
Critical device issues	*1%*	*2%*	*1%*	*-1%*	Improvement over interval
-No workstation functionality (e.g., does not boot up)	*1*	*2*	*0*		
-User interface non-functional (e.g., damaged mouse / keyboard)	*0*	*0*	*1*		
End-user interface (based on workstation location and provider roles)					
-Tap-in/-out logon/logoff device present and functional	84%	84%	67%	-16%	Decline over interval
-Fingerprint-based authentication device present and functional	n/a#	15%	6%	-9%	Decline over interval; superseded by smartphone 2-factor authentication
-Real-time dictation transcription device present and functional	73%	98%	100%	2%	Improvement over interval with implementation of smartphone-based mobile dictation system
Electronic medical record printing (*e.g.,* discharge instructions) functional	78%	95%	94%	-1%	Decline over interval
System security and maintenance applications active	100%	100%	100%	0%	System-wide push security updates and automated monitoring
Mobile workstations	n=27	n=29	n=26		
Critical device issues	*15%*	*24%*	*4%*	*-20%*	Improvement over interval
-No workstation functionality (e.g., does not boot up)	*3*	*4*	*0*		
-No network connectivity	*1*	*1*	*0*		
-Workstation battery not holding charge	*0*	*2*	*0*		
-Excessively slow performance	*0*	*0*	*1*		
End-user interface (based on workstation location and provider roles)					
-Tap-in/-out logon/logoff device present and functional	57%	52%	62%	10%	Improvement over interval
-Fingerprint-based authentication device present and functional	n/a#	0%	5%	5%	Improvement over interval; superseded by smartphone 2-factor authentication
-e-signature capture device present and functional	74%	73%	85%	12%	Improvement over interval
-Real-time dictation transcription device present and functional	n/a	n/a	100%	100%	Improvement over interval with implementation of smartphone-based mobile dictation system
Electronic medical record printing (*e.g.,* discharge instructions) functional	n/a#	n/a#	88%	n/a	New functionality implemented during interval
System security and maintenance applications active	100%	100%	100%	0%	System-wide push security updates and automated monitoring

Not implemented at time of inventory.

Table 2: ED provider end-user anonymous survey results.

ED Clinical Personnel Web Survey	April 2016 (Pre-intervention baseline)	April 2018 (Post-intervention follow-up)
Respondent Characteristics	22% (108 of 495 ED personnel)	44% (219 of 501 ED personnel)
Primary provider role		
-Advanced practice provider	5%	6%
-Certified nursing assistant	6%	3%
-Physician	40%	40%
-Registered Nurse	39%	42%
-Scribe	6%	5%
-Other	5%	6%
Age (years; mean ± SD)	39 ± 10 (range 22-66)	39 ± 11 (range 22-68)
Sex		
-Female	53%	56%
-Male	45%	43%
-Other	0%	0%
-Prefer not to answer	2%	2%
Comfort level with information technologies in general (1="very uncomfortable"; 5="very comfortable"; median ± IQR1-3)	4 (4-5)	4 (4-5)

ED CIS Web Survey Responses (1="very poorly / inconsistently"; 4="very well / consistently"; median ± IQR1-3 [no. of respondents§])	April 2016 (Pre-intervention baseline)	April 2018 (Post-intervention follow-up)	p (Mann-Whitney 2-way; significance set at α=0.05)
Perception of how well ED CIS usually works as intended for:			
-General system functions	3 (2-3; n=92)	2 (2-3; n=195)	0.099
-Patient registration	3 (3-3; n=73)	3 (3-3; n=153)	0.368
-Chart access and order entry	3 (3-3; n=90)	3 (3-3; n=191)	0.638
-Dictation functions	2 (1-3; n=43)	3 (2-3; n=86)	**0.005**
-Laboratory results access and review	3 (3-3; n=92)	3 (3-3; n=189)	0.704
-Imaging results access and review	3 (3-3; n=87)	3 (3-3; n=180)	0.818
-Printing	2 (1-2; n=93)	2 (1-2; n=195)	0.984
-Web access (Intranet, Web email)	3 (2-3; n=93)	3 (2-3; n=195)	0.072
-Web access (search, *e.g.*, Google)	3 (2-3; n=92)	3 (2-3; n=194)	0.347
Proportion of your shifts in last month that were hindered by an ED CIS issue that significantly impacted your ability to complete clinical duties (%; mean ± SD)	25 ± 27 % (range 0-100%)	29 ± 28 % (range 0-100%)	0.124
Proportion of encountered ED CIS issues for which you started an I.S. work-ticket (%; mean ± SD)	13 ± 2 % (range 0-100%)	20 ± 28 % (range 0-100%)	**0.028**
End-user process to start an ED CIS repair request:			
-Call I.S. support line	47%	45%	
-Contact unit secretary	31%	33%	
-Contact supervisor	2%	2%	
-Use Web portal to create user-generated work ticket	n/a#	2%	n/a
-I have never submitted a repair request	8%	5%	
-I do not know how to submit a work ticket	1%	2%	
-I do not have the time to submit a repair request	9%	8%	
-Other	2%	2%	
What proportion of ED CIS issues are corrected when you do not submit a work-ticket? (%; mean ± SD)	25 ± 30 % (range 0-100%)	24 ± 30 % (range 0-100%)	0.787
What proportion of ED CIS issues are corrected when you do submit a work-ticket	58 ± 38 % (range 0-100%)	62 ± 36 % (range 0-100%)	0.555

Key: CIS = Clinical Information Services; ED = Emergency Department; IQR = interquartile range; I.S. = Information and Network Services; SD = standard deviation
§ The number of responses for each CIS function category was determined by its relevance to respondents' clinical provider roles.
# Not implemented at time of first survey.

Table 3: Characteristics of ED work-tickets submitted.

	Historical Data					Pre-Intervention Baseline Assessment Period	Intervention Period				Post-Intervention Assessment Period	Total
	9/2014-12/2014	1/2015-4/2015	5/2015-8/2015	9/2015-12/2015	1/2016-4/2016	5/2016-8/2016	9/2016-12/2016	1/2017-4/2017	5/2017-8/2017	9/2017-12/2017	1/2018-4/2018	
Total ED work-tickets	33	72	43	63	245	273	320	232	175	268	338	2,062
Submitted through new online portal	n/a	n/a	n/a	n/a	n/a	n/a	22 (7%)	14 (6%)	14 (8%)	6 (2%)	20 (6%)	76 (4%)
Not ED CIS	14	7	10	11	86	133	147	93	78	115	181	875
Total ED CIS-related work-tickets	19 (58%)	65 (90%)	33 (77%)	52 (82%)	159 (65%)	140 (51%)	173 (54%)	139 (60%)	97 (55%)	153 (57%)	157 (46%)	1,187 (58%)
General system functions	26%	37%	31%	36%	26%	29%	29%	31%	41%	22%	37%	32%
Patient registration / demographics	0%	6%	6%	10%	10%	8%	14%	10%	6%	2%	6%	8%
Access to previous EMR chart/records	0%	3%	0%	0%	1%	1%	1%	1%	1%	3%	1%	1%
Imaging results access + review	0%	2%	3%	0%	3%	3%	12%	8%	6%	3%	5%	5%
Laboratory results access + review	21%	0%	3%	0%	0%	1%	1%	0%	0%	1%	1%	1%
Patient disposition (admitting / discharge / transfer)	0%	0%	0%	2%	1%	3%	1%	0%	1%	1%	1%	1%
Dictation functions	0%	3%	9%	2%	1%	3%	0%	1%	0%	0%	1%	1%
Printing	32%	23%	33%	25%	25%	20%	17%	16%	13%	13%	13%	18%
Inter-/Intra-net access	21%	26%	15%	25%	33%	32%	25%	33%	32%	55%	35%	33%
Time to work-ticket completion (days; mean ± SD)	2.7±5.9	4.6±9.0	3.3±7.6	10.7±4.5	9.9±2.1	7.8±2.3	7.2±1.4	7.4±2.6	7.1±1.2	7.2±1.1	7.1±1.6	7.5±1.8

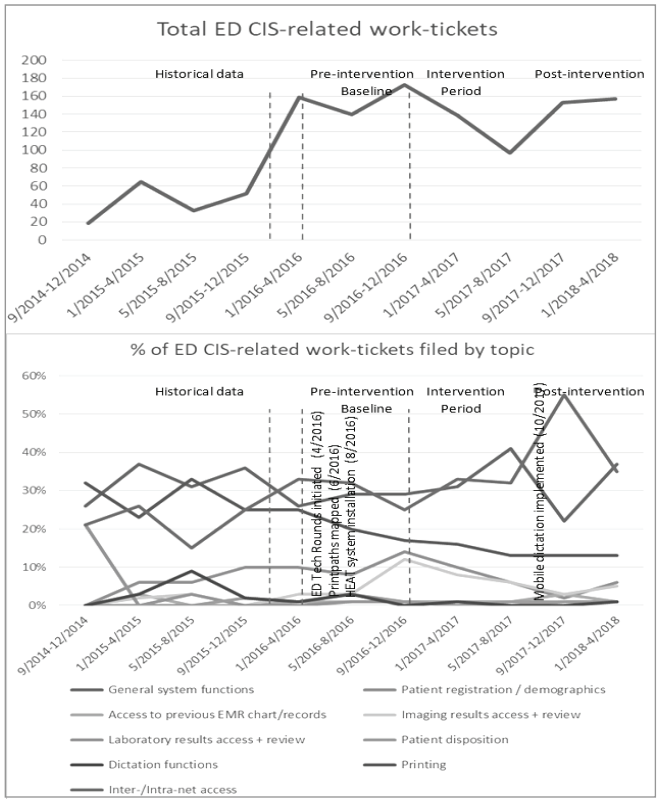


Figure 2: Trends in number and topic of work-tickets filed over time.

A. ED CIS Infrastructure Operational Knowledgebase

The project team conducted a structured physical inventory, integrative mapping, and function and functionality baseline assessment of CIS in the project ED. This process collected the core information necessary to ensure, monitor, and sustain critical ED CIS functions, along with a structured checklist process with metrics and an assessment of common failure points. As a result, high-impact device problems that impeded the performance of clinical tasks were identified along with additional non-critical equipment and function issues– these had not been reported and addressed through the existing service request process. Furthermore, only 13% of Web survey respondents reported initiating service requests when they encountered ED CIS issues, a stark disconnect given the large number of providers who indicated that CIS problems had significantly impacted their ability to complete clinical duties during the preceding month. The project's Web survey provided some insight into this discrepancy, in that 10% of respondents expressed that they did not have time / know how to submit repair requests. Conversely, several respondents indicated that they preferred to attempt to fix CIS devices themselves.

In parallel with the barriers that prevented consistent problem identification, reporting, and repair, the absence of a proactive ED CIS infrastructure maintenance and sustainment mechanism was identified. As potentially unifying failure points that underpinned recurring ED CIS issues, these findings were presented to the institutional I.S. leadership for discussion. All project-developed materials were provided as a comprehensive (snap-shot) accounting of clinically-relevant ED CIS infrastructure, their expected functions, and identified problems. Of note, the reference materials currently remain in use for strategic and operational planning. (As proof of value, these materials helped reduce the time required by the project team to complete an ED CIS inventory assessment from 40 hours in 2016 to 24 hours in 2018.)

B. Targeted Interventions and Post-intervention Follow-up

The QI interventions derived from the baseline ED CIS reassessment were implemented with variable success and varying results based on findings from the 2018 re-assessment and survey:

1. Creation of the inventory-map document with uniform labeling, specification and confirmation of expected functions, and highlighting of active issues had immediate utility for I.S. teams conducting repairs, although follow-up testing did not reveal consistent, concrete evidence of improvements directly attributable to the interventions.

2. With prioritization of the repair / replacement of malfunctioning / absent ED CIS devices, fewer critical device issues (especially in mobile workstations) were found at follow-up. The institution-driven deployment of mobile app-based solutions for user authentication and dictation functions increased provider-level implementations of core clinical ED CIS functions.

3. The successful implementation of ED "Tech Rounds" was followed by service interruption then resumption at lower frequency– its effects were difficult to gauge. Anecdotal reports indicated that the assigned administrative assistant remained constructively busy for the allotted time, supporting the need for sustainment of this role.

4. Several months after introduction of a self-service work-ticket Web portal, clinical providers were using it to submit a small proportion of new I.S. work-tickets. Clear patterns in work-tickets submitted were not detected.

VI. Conclusion

The project's initial and follow-up assessments of clinically-relevant ED CIS infrastructure, their expected functions, and identified problems demonstrated how the complex intersection of health care and informatics presents dynamic challenges that can impede patient care. Furthermore, the continuous advances in CIS and HIT were noted to introduce new problems, compound existing ones, and undermine developed solutions. The need for proactive planning, ongoing vigilance, and multi-disciplinary collaboration with meaningful support for ED CIS operational sustainment is clear.

VII. Acknowledgments and Competing Interests

Acknowledgments

AG, JT, GJ, LK conceived the project, designed the trial, and obtained research funding. AG, JT, DP, CP, LK supervised the conduct of the trial and data collection. AG, JT, AD, RH, MJ, SS, AS undertook data acquisition and managed the data. AG, JT, DP, CP, GJ, LK analyzed the data; AG, JT, LK drafted the manuscript, and all authors contributed substantially to its preparation. AG, JT, LK take responsibility for the paper as a whole. Please contact Amelia L. Gurley (amelia_gurley@brown.edu) with any questions or further communications.

Competing Interests

Excerpts from the manuscript were presented at the 2016 IEEE International Conference on Healthcare Informatics in Chicago, IL, and the 2018 Institute for Healthcare Improvement Scientific Symposium in Orlando, FL. The manuscript is based on work supported by the Depart-ment of Emergency Medicine at Alpert Medical School of Brown University and Rhode Island Hospital (RIH). The authors have no relevant financial information or potential conflicts to disclose. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the Department or RIH.

VIII. References

[1] C. A. Pedersen, P. J. Schneider, and D. J. Scheckelhoff. ASHP national survey of pharmacy practice in hospital settings: prescribing and transcribing- 2016. Am J Health Syst Pharm., vol. 74(17), pp. 1336-1352, 2017.

[2] U.S. Department of Health and Human Services. Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009. Available at: https://www.healthit.gov/sites/default/files/hitech_act_excerpt_from_arra_with_index.pdf

[3] J. S. Ash, M. Berg, and E. Coiera. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc., vol. 11(2), pp. 104-112, 2004.

[4] E. M. Campbell, D. F. Sittig, J. S. Ash, et al. Types of unintended consequences related to computerized provider order entry. J Am Med Inform Assoc, vol. 13(5), pp. 547-556, 2006.

[5] R. B. Myers, S. L. Jones, D. F. Sittig, et al. Review of reported clinical information system adverse events in US Food and Drug Administration databases. Appl Clin Inform, vol. 2(1), pp. 63-74, 2011.

[6] E. Joffe, C. F. Bearden, M. J. Byrne, et al. Duplicate patient records—implication for missed laboratory results. AMIA Annu Symp Proc 2012, pp. 1269-1275.

[7] J. S. Adelman JS, G. E. Kalkut, C. B. Schechter, et al. Understanding and preventing wrong-patient electronic orders: a randomized controlled trial. J Am Med Inform Assoc, vol. 20(2), pp. 305-310, 2013.

[8] D. W. Meeks, M. W. Smith, L. Taylor, et al. An analysis of electronic health record-related patient safety concerns. J Am Med Inform Assoc., vol. 21(6), pp. 1053-1059, 2014.

[9] D. F. Sittig, D. Gonzalez, and H. Singh. Contingency planning for electronic health record-based care continuity: a survey of recommended practices. Int J Med Inform, vol. 83(11), pp. 797-804, 2014.

[10] B. H. Just, D. Marc, M. Munns, et al. Why patient matching is a challenge: research on master patient index (MPI) data discrepancies in key identifying fields. Perspect Health Inf Manag, vol. 13(1e), 2016.

[11] D. F. Sittig, A. Wright, J. Ash, et al. New unintended adverse consequences of electronic health records. Yearb Med Inform. vol. 10(1), pp. 7-12, 2016.

[12] Y. Wang, E. Coiera, B. Gallego, et al. Measuring the effects of computer downtime on hospital pathology processes. J Biomed Inform., vol. 59, pp. 308-315, 2016.

[13] A. Wright, T. T. Hickman, D. McEvoy, et al. Analysis of clinical decision support system malfunctions: a case series and survey. J Am Med Inform Assoc., vol. 23(6), pp. 1068-1076, 2016.

[14] D. Lyell, E. Coiera. Automation bias and verification complexity: a systematic review. J Am Med Inform Assoc., vol. 24(2), pp. 423-431, 2017.

[15] D. Lyell, F. Magrabi, M. Z. Raban, et al. Automation bias in electronic prescribing. BMC Med Inform Decis Mak., vol. 17(1), p. 28, 2017.

[16] E. Larsen, A. Fong, C. Wernz, et al. Implications of electronic health record downtime: an analysis of patient safety event reports. J Am Med Inform Assoc., vol. 25, pp. 187-191, 2018.

[17] D. F. Sittig, A. Wright, E. Coiera, et al. Current challenges in health information technology-related patient safety. Health Informatics J., Dec 11, 2018:1460458218814893.

[18] J. R. Nebeker, J. M. Hoffman, and C. R. Weir. High rates of adverse drug events in a highly computerized hospital. Arch Intern Med., vol. 165(10), pp. 1111-1116, 2005.

[19] J. I. Westbrook, M. T. Baysari, L. Li, et al. The safety of electronic prescribing: manifestations, mechanisms, and rates of system-related errors associated with two commercial systems in hospitals. J Am Med Inform Assoc., vol. 20(6), pp. 1159-1167, 2013.

[20] M. G. Amato, A. Salazar, T. T. Hickman, et al. Computerized prescriber order entry-related patient safety reports: analysis of 2522 medication errors. J Am Med Inform Assoc. vol. 24(2), pp. 316-322, 2017.

[21] M. O. Kim, E. Coiera, F. Magrabi. Problems with health information technology and their effects on care delivery and patient outcomes: a systematic review. J Am Med Inform Assoc., vol. 24(2), pp. 246-250, 2017.

[22] J. L. Howe, K. T. Adams, A. Z. Hettinger, et al. Electronic health record usability issues and potential contribution to patient harm. J Am Med Assoc., vol. 319(12), pp. 1276-1278, 2018.

[23] M. L. Graber, D. Siegal, and H. Riah. Electronic health record-related events in medical malpractice claims. J Patient Saf., vol. 15(2), pp. 77-85, 2019.

[24] F. Magrabi, M. S. Ong, W. Runciman, et al. Patient safety problems associated with healthcare information technology: an analysis of adverse events reported to the US food and drug administration. AMIA Annu Symp Proc. 2011, pp. 853-857.

[25] F. Magrabi, M. Baker, I. Sinha, et al. Clinical safety of England's national programme for IT: a retrospective analysis of all reported safety events 2005 to 2011. Int J Med Inform. vol. 84(3), pp. 198-206, 2015.

[26] J. L. Wiler, C. Gentle, J. M. Halfpenny, et al. Optimizing emergency department front-end operations. Ann Emerg Med. vol. 55(2), pp. 142-160, 2010.

[27] L. I. Horwitz, T. Meredith, J. D. Schuur, et al. Dropping the baton: a qualitative analysis of failures during the transition from emergency department to inpatient care. Ann Emerg Med., vol. 53(6), pp. 701-710, 2009.

[28] L. Poissant, J. Pereira, R. Tamblyn, et al. The impact of electronic health records on time efficiency of physicians and nurses: a systematic review. J Am Med Inform Assoc, vol. 12(5), pp. 505-516, 2005.

[29] S. Spellman Kennebeck, N. Timm, M. K. Farrell, et al. Impact of electronic health record implementation on patient flow metrics in a pediatric emergency department. J Am Med Inform Assoc., vol. 19(3), pp. 443-447, 2011.

[30] D. J. Mathison, and J. Chamberlain. Evaluating the impact of the electronic health record on patient flow in a pediatric emergency department. Appl Clin Inform., vol.2(1), pp. 39-49, 2011.

[31] D. A. Propp. Successful introduction of an emergency department electronic health record. West J Emerg Med., vol. 13(4), pp. 358-361, 2012.

[32] S. Syed, D. Wang, D. Goulard, et al. Computer order entry systems in the emergency department significantly reduce the time to medication delivery for high acuity patients. Int J Emerg Med., vol. 6(1), pp. 20, 2013.

[33] J. J. Brixey, D. J. Robinson, Z. Tang, et al. Interruptions in workflow for RNs in a level one trauma center. AMIA Annual Symposium Proceedings 2005, pp. 86-90.

[34] J. D. D'Amore, J. C. Mandel, D. A. Kreda, et al. Are meaningful use state 2 certified EHRs ready for interoperability? Findings from the SMART C-CDA Collaborative. J Am Med Inform Assoc., vol. 21(6), pp. 1060-1068, 2014.

[35] A. J. Holgren, and E. W. Ford. Assessing the impact of health system organizational structure on hospital electronic data sharing. J Am Med Inform Assoc., vol. 25(9), pp. 1147-1152, 2018.

[36] H. Hirano H. "5 pillars of the visual workplace- The sourcebook for 5S implementation." New York: Productivity Press, 1995.

Development of an algorithm for optimizing the technological process of kaolin enrichment

Ibrokhimali Normatov
Department of "Applied mathematics and intellectual technologies
National University of Uzbekistan named after Mirzo Ulugbek
Tashkent, Uzbekistan
ibragim_normatov@mail.ru

Eshmurod Kamolov
Chirchik State Pedagogical Institute of Tashkent
Amir Temur street, 104 111700
Tashkent, Uzbekistan
e-mail: kamolov_eshmurod@mail.ru

***Abstract*— An algorithm has been developed to optimize the technological process of kaolin enrichment, based on the fuzzy sets theory; it allows evaluating the objects according to the grade of membership in a certain class and the gradient method. The dependences of the gradient values and the redistribution of the weights on initial values, as well as on the degree of correlation of the distributions of these weights, are considered in the paper.**

Keywords—kaolin, kaolin enrichment, biotechnological system, technological process, maximization of productivity, negatively correlated weights.

I. INTRODUCTION

In the process of implementing the management system of the biotechnological system (BTS) one has to repeatedly come up against the choice of optimization method from several alternatives. The construction of a relevance tree [1] with an indication of relative priorities becomes insufficient, since it is necessary to take into account the subjective character of optimization tools evaluation and the redistribution of priorities as the system develops. There is a need to develop such methods of solution substantiation, which to a greater extent would meet the conditions of optimization systems. When comparing alternative solutions, there is an aim to bring dissimilar goals to a single basis - the degree of achievement.

II. MATERIAL AND RESEARCH METHODS.

Solving the technological process (TP) optimization problem using economic efficiency indices as a criterion for optimality is a difficult task, since it is necessary to take into account the simultaneous influence of a large number of variables on which the constraints of various types accumulate. The generalized optimality criterion R of economic efficiency of a (TP) is a certain function of the following indices:

$$R = f(C,\ K,\ D,\ P,\ U, B)$$

Consider these figures in relation to the process of kaolin enrichment.

The net cost of produce C as an optimality criterion has significant drawbacks, namely, this index does not take into account the output quality of the product K and has a low sensitivity to control actions. Therefore, to apply the criterion $C \to \min$ to control the object under consideration is impractical.

The role of the quality index $K \to \max$ of the output product becomes clear only in pricing. Obviously, the price of the product of higher quality should be higher. Therefore, at present such indices as profit P and income D obtained from sales of products are widely used as a criterion for maximality.

However, the use of the indices $P \to \max$ and $D \to \max$ in the direct form is possible only for commercial products that have a certain known price. For intermediate products (such as kaolin) conditional prices should be applied.

The conditional profit index U in various cases degenerates into simpler indices. So, if the productivity B of the stage is set then the maximization task $P \to \max$ becomes equivalent to minimization of the net cost

$$C \to \max:\ P \to \max = C \to \max .$$

A similar result is obtained when the difference between them and the net cost of production is relatively small. If the price is significantly higher than the net cost, then the maximization of the conditional profit $U \to \max$ is equivalent to the maximization of productivity

$$B \to \max : U \to \max = B_{mp} \to \max$$

Taking into account there quirement so simplicity, completeness, equivalence, universality, include in economic (maximization of productivity) and technological (maximization of the target product) indices, we obtain:

$$B \to max\ = B_{mp} \to max$$

So, as a criterion for optimality we will use such an index as maximization of the productivity of a given planned net cost:

$$B \to \max,\ C = const = C_{pnc},$$

or a criterion of maximum output of the target product.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

The fact that the relationship between the degree of aim achievement and the quality indices is not strictly defined and does not create insurmountable difficulties in the analysis, since the modern theory of fuzzy sets [1, 2] allows evaluating the objects according to the grade of membership in a certain class. For each parameter of the technological process of kaolin enrichment, the range can be specified in which the degree of its compliance with the specified goals (μ) varies from 0 to 1 or the error from the inconsistency of the actual value of the parameter to the specified one $(\bar{\mu})$ varies from 1 to 0. The dependence of the error on the value of each parameter, for example, the calculation rate or the modeling error, can be approximated by an analytical expression of the form

$$\overline{\mu_i} = [1 + a(\alpha_{oi} - \alpha)_i^2]^{-1}, \alpha_i \le \alpha_\alpha, \quad (1)$$

where α_{oi} and α_i are initial and current value soft the parameter.

The quality of the system, characterized by several indices, at a fuzzy definition of their values is expressed by the intersection function of fuzzy sets

$$\overline{\mu_\Sigma} = \max_i(\overline{\mu_i}), \quad (2)$$

however, this criterion cannot be applied to the assessment of components of the management system, since it does not fix the improvement in systemquality when one or several indices change, the errors of which are not the maximum ones. In the absence of information about the influence of individual system parameters on its quality, a multiple criterion is used; its components have weights determined by expert judgment

$$\overline{\mu_\Sigma} = \sum_i \omega_i \overline{\mu_i}\,. \quad (3)$$

This criterion, however, does not take into account the dependence of the weights on the value of individual indices [2, 3]. It would be more expedient to redistribute the weights in proportion to the fraction of the weighted error

$$\omega_i' = \frac{\omega_i \overline{\mu_i}}{\sum \omega_k \overline{\mu_k}}, \quad (4)$$

but in this case, every change in the *i*-th index, leading to a change in the general criterion (3), causes a redistribution not only of ω_i, but also of $\overline{\mu_i}$ to maintain the achieved value of $\overline{\mu_\Sigma}$. A new set of values $\overline{\mu_i}$ requires a new redistribution of the weights, and this recursion has no convergence.

The problem of constructing a multiple criterion is greatly simplified if we take into account the fact that information about the weights of the indices is needed to make only one decision, and the level of quality of the system achieved at this point of time does not affect the decision. So, it is enough to assume that the weights of the indices are determined according to (4) from the previous step, and the error values are conventionally taken equal to one.

The scale of the error function is changed, but the derivatives in terms of the quality indices of the system stay unchanged. It is significant that such an approach allows for a more rigorous consideration of the second component of the system quality criterion - the cost on improving its indices.

Assume that at the initial stage of work, the costs are distributed unevenly across indices, i.e. have weight ϑ_i. In this case, the quality multiple criterion will be

$$M = a_1 \sum \omega_i \overline{\mu_i} + a_2 \sum \vartheta_i (1 - \overline{\mu_i}), \quad (5)$$

The augend characterizes the error from the inconsistency of the indices of the system with the specified values, the addend - the costs of eliminating this discrepancy. As a result of the work, a certain effect must be achieved, which can be measured in fractions of the initial error value: [4, 5]

$$M(\overline{\mu_i} = 1) = a_1;\ M(\overline{\mu_i} = 0) = a_2;$$

$$\Delta M = a_1(1 - \frac{a_2}{a_1});\ a_1 + a_2 = 1.$$

Equation (5) may be written in the form:

$$M = a_2 + \sum (a_1\omega_1 - a_2\vartheta_i)\overline{\mu_i}, \quad (6)$$

here a_1 and a_2 are the dimensionless coefficients, the ratio of which determines the efficiency of the work performed. Since criterion (6) is designed to compare alternatives that reduce the error for each index by the value $\Delta\overline{\mu_i}$, the coefficients a_1 and a_2 should be considered as derivatives of the resulting error and costs for the weighted average components. Assuming that each solution in the optimization area is efficient, i.e. $\Delta M > 0$, coefficients a_1 and a_2 can be obtained by expert evaluation, based on the specific situation in which alternatives are considered, and on the relative complexity of the decision. Obviously, at given a_1 and a_2, the preference should be given to a decision that minimizes the criterion *M*. However, in the Pareto-optimal region, a deeper analysis of alternatives is required, taking into account, in particular, the uneven distribution of costs between the error components. Despite the complexity of the comparative assessment of costs for each index in a comprehensive solution, it becomes reasonable to take into account the weights ϑ_i, bearing in mind that the costs of various resources are not directly

aimed at changing $\overline{\mu_i}$, but at changing the system quality indices α_i related to dependence $\overline{\mu_i}$ (1).

III. DISCUSSION OF RESULTS.

Assume that costs are related to the changes in system indices by expression

$$\Delta Z_i = Z'_{\alpha i}\Delta\alpha, \quad (7)$$

moreover, as the system develops, the unit costs are reduced, as there is an accumulation of its functional capabilities. The change in errors as the index changes is a derivative $\mu'_{\alpha i}$. Then the following relations are true:

$$\Delta Z_i = \frac{Z'_{\alpha i}}{\mu'_{\alpha i}}\Delta\overline{\mu_i} = r_i\Delta\overline{\mu_i}, \quad (8)$$

$$\Delta Z\Big|_{\Delta\overline{\mu}=1} = \sum_k r_k, \quad (9)$$

$$\Delta M = a_1\sum\omega_i\Delta\overline{\mu_i} - a_2\sum\vartheta_i\Delta\overline{\mu_i}, \quad (10)$$

$$\vartheta_i = r_i : \sum_k r_k. \quad (11)$$

Here $\Delta\mu_i$ is defined as the difference between the error values before and after the next step, counting from the initial state of the system. The change in conditions may require correction of the costs dependence on error. It is obvious that a change in the error components should ensure motion along the normal to a hypersurface of equal level *M*. If several options give the same value of *M*, the option that leads to the maximum value of the *M* gradient at the next decision step is chosen [6, 7].

Equation of the normal at a point $\overline{\mu_i}=1,\ i=\overline{1,n}$ has the form:

$$\frac{1-\overline{\mu_1}}{a_1\omega_1 - a_2\vartheta_1} = \frac{1-\overline{\mu_2}}{a_1\omega_2 - a_2\vartheta_2} = \ldots = C. \quad (12)$$

The value of the criterion gradient is

$$|gradM| = \sqrt{\sum_{i=1}^{n}(a_1\omega_i - a_2\vartheta_i)^2} = \\ = \sqrt{n[a_1^2 D(\omega) - 2a_1a_2\operatorname{cov}(\omega,\vartheta) + a_2^2 D(\vartheta)] + \frac{1}{n}(a_1-a_2)^2} \quad (13)$$

where $D(\omega),\ D(\vartheta)$ are the variance of the weights of the errorcomponents and the costs of changing the parameters of the system; $\operatorname{cov}(\omega,\vartheta)$ is the covariance of weights ω_i and ϑ_i.

Neglecting the constant components, consider the criterion

$$K = D(\omega) - 2\frac{a_2}{a_1}\operatorname{cov}(\omega,\vartheta) + (\frac{a_2}{a_1})^2 D(\vartheta). \quad (14)$$

When changing the components of the errors $\overline{\mu_i}$, their weights, according to the accepted assumptions, change as follows:

$$\omega'_i = \frac{\omega_i\overline{\mu_i}}{\sum_k \omega_k\overline{\mu_k}}. \quad (15)$$

When moving in the direction of the normal

$$\omega'_i = \omega_i\frac{1-c(a_1\omega_i - a_2\vartheta_i)}{1-c(a_1\sum_k\omega_k^2 - a_2\sum_k\omega_k\vartheta_k)}. \quad (16)$$

Consider the dependence of the magnitude of the gradient *M* and the redistribution of the weights ω_i on the initial values of ω_i and ϑ_i, as well as on the degree of correlation of the distributions of these weights.

The weights of the cost components are:

$$\vartheta_i = \frac{1}{n},. \quad (17)$$

here

$$D(\vartheta) = \operatorname{cov}(\omega,\vartheta) = 0, \quad 18)$$

$$K_1 = D(\omega) = \sigma^2(\omega), \quad (19)$$

where σ is the standard deviation of the weights.

Introduce the coefficient *P* which characterizes the difference in weights ω_i and ω'_i:

$$\omega'_i = \omega_i(1+P_i), \quad (20)$$

Then according to (12),

$$P_i = \frac{1-c(a_1\omega_i - a_2\vartheta_i) - 1 + c(a_1\sum\omega_k^2 - a_2\sum\omega_k\vartheta_k)}{1-c(a_1\sum\omega_k^2 - a_2\sum\omega_k\vartheta_k)} \cong \quad (21)$$
$$\cong ca_1(\sum\omega_k^2 - \omega_i) + ca_2(\vartheta_i - \sum\omega_k\vartheta_k)$$

Under condition (17) $\sum\omega_k\vartheta_k = \frac{1}{n}$,

$$P_i = ca_1(\sum\omega_k^2 - \omega_i) = ca_1\left[nD(\omega) - \Delta\omega_i\right], \quad (22)$$

where $\Delta\omega_i = \omega_i - \frac{1}{n}$.

According to (22), small weights increase, large ones decrease; this leads to a decrease in dispersion. From (19) it is seen that, the change in error slows down at further motion along the normal.

IV. CONCLUSIONS

In the case under consideration, when moving along the normal, the difference in weights of certain indices and the rate of change of the resulting damage decreases, i.e. the approaching speed of the optimized system indices to the specified onesslows down. The fact, that negatively correlated weights ω_i and ϑ_i give the maximum value of the gradient is consistent with the intuitive idea of the optimality of the conditions when the costs of achieving the most important goals are minimal. Since the dependence of costs on system quality indices is non-linear, the optimal strategy for reducing the resulting error should differ in the sections

$$1 \geq \overline{\mu_i} \geq 0,5 \text{ and } 0,5 \geq \overline{\mu_i} \geq 0,$$

that is, at the initial and final stages of solving the problem.

The advantage of such a decision-making procedure is the possibility of an account of any changes in external conditions in the process of multi-stage work; for example, the relative value of various system parameters, the error dependence on quality indices, and the costs allocation for improving various indices.

On the other hand, the focus on maintaining a high rate of reduction of the resulting error is especially relevant for creating complex human-machine systems with limited resources, since this strategy allows the user to be interested in work continuation.

Based on the above methods, an algorithm for optimizing the process of kaolin enrichment has been developed. Figure 1. a) and b) shows the results of optimization and Fig.2. a) and b) shows maximum rate of iron, different at an initial pH 2.2-2.3.

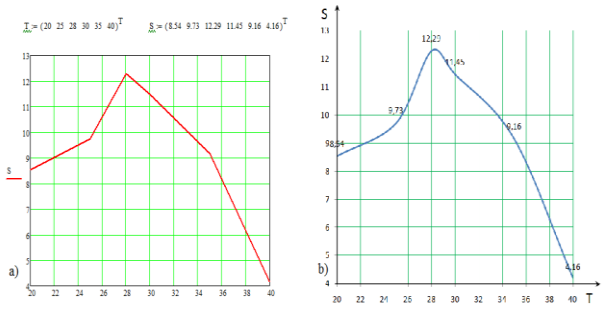


Fig.1. Maximum rate of iron oxidation at reference temperature of 28-30

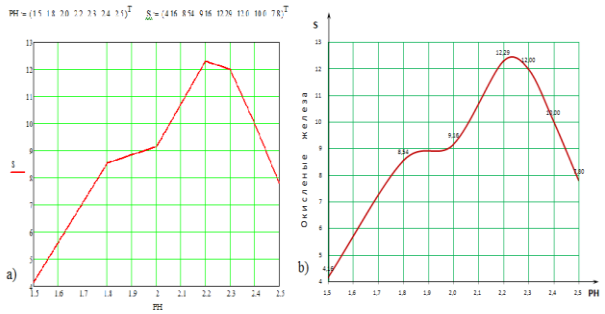


Fig. 2. Maximum rate of iron, different at an initial pH 2.2-2.3

TABLE 1. ENRICHMENT OF THE COMPOSITION OF ORES OF GRAY KAOLIN DEPENDING ON TIME-1

t (time)	SiO_2 % Silica	TiO_2 % Titanium	Al_2O_3 % Aluminium	Fe_2O_3 % Iron	K_2O % Potassium
6	53,5	0,6	27,8	2,8	1,7
12	51,8	0,6	32,1	1,6	1,6
18	50,2	0,6	34,1	0,9	1,6
24	49,8	0,6	34,8	0,9	1,6
30	48,9	0,7	35,1	0,8	1,6

TABLE 2. ENRICHMENT OF THE COMPOSITION OF ORES OF GRAY KAOLIN DEPENDING ON TIME-2

$t\ (time)_2$	SiO_2 Silica	TiO_2 Titanium	Al_2O_3 Aluminium	Fe_2O_3 Iron	K_2O Potassium
6	5,351	0,0063	0,2780	0,0277	0,0168
12	0,5178	0,0064	0,3204	0,0158	0,016
18	0,518	0,0066	0,3413	0,0098	0,0164
24	0,4978	0,0065	0,3480	0,0090	0,0162
30	0,4890	0,0067	0,3510	0,0080	0,0163

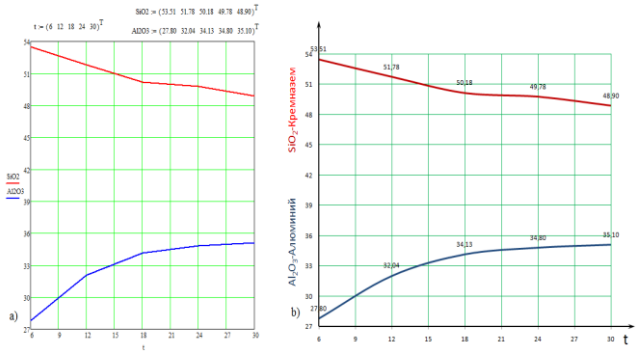


Fig. 3. Changes in chemical composition of ores in the process of kaolin enrichment

REFERENCES

[1] A.V.Kabulov, I.H.Normatov, A.O.Ashurov, "Computational methods of minimization of multiple functions Journal of Physics", 2019, p.1260.

[2] A.V.Kabulov, I.H.Normatov, "About problems of decoding and searching for the maximum upper zero of discrete monotone functions", Journal of Physics, 2019, p.1260.

[3] I.H.Normatov, "Principle of independence of continuation of functions multivalued logic from coding", Journal of Physics, 2018, p.210.

[4] A.V.Kabulov, I.H.Normatov, A.Karimov, "Algorithmization control of complex systems based on functioning tables", Journal of Physics, 2020, p.1441.

[5] Ch.J.Herget, "Automated design of control systems", Mechanical engineering, Moscow,1989.

[6] D.A.Pospelov, "Fuzzy sets in models of control and artificial intelligence", Science, Moscow, 1986.

[7] V.M.Kantere, "Use of mathematical methods in microbiology", Pushino,1975, p.26.

[8] A.B.Rubin, N.F.Pyteva, T.Yu.Reznichenko, "Kinetics of biological processes", Moscow State University, Moscow,1977.

[9] M.I.Schadov, K.E.Vinnitsky, "Development of engineering and technology of open coal mining", Nedra, Moscow, 1987.

[10] V.S.Balakirev, V.M.Volodin, A.M.Tsirlin, "Optimal control of chemical technology processes", Chemistry, Moscow, 1978.

[11] "Overview of the market of raw kaolin and enriched kaolin in the CIS countries", 9th edition, 25.03.2019, p.179.

Feature Selection for Deep Neural Networks in Cyber Security Applications

Alexander Davis
California State University, Fresno
Fresno, CA
alexanderdvs@mail.fresnostate.edu

Sumanjit Gill
California State University, Fresno
Fresno, CA
sumanjitgill@mail.fresnostate.edu

Robert Wong
California State University, Fresno
Fresno, CA
robertwong@mail.fresnostate.edu

Shahab Tayeb
California State University, Fresno
Fresno, CA
tayeb@mail.fresnostate.edu

***Abstract*—To stay ahead of novel attacks, cybersecurity professionals are developing new software and systems using machine learning techniques. Neural network architectures improve such systems, e.g., IDS, by implementing anomaly detection, which differentiates benign packets from malicious packets. In order for an IDS to best predict anomalies, the data the model is trained on is typically pre-processed through normalization and feature selection/reduction. The impacts of pre-processing techniques play an important role in training a neural network to optimize its performance. This paper proposes a DNN with 2 hidden layers for the IDS architecture and compares two commonly used normalization pre-processing techniques. We find that using ZScore normalization increased performance in accuracy by 4.46% reduced loss by 2.04%, improved F-Score by 4.70%, and increased AUC-ROC by 23.64%.**

***Index Terms*—**

I. Introduction

Research on machine learning and neural networks in intrusion detection systems (IDS) has become more prominent as the body of knowledge on neural networks has expanded [1]. IDS have been a major topic of research in the network security field due to the importance of identifying intrusions, attacks, and violations, but have also been criticized for improvement because of its importance to security [2] [3]. IDS are vital to network security since they aim to prevent disruptions in network communications [4]. IDS typically use a statistic probability model, machine learning, or a neural network to determine if data is malicious; neural networks, however, have the capability of identifying deeper characteristics from traffic [5]. IDS offer an improvement in security by predicting and preventing malicious attacks using machine learning [6] [7]. Previously, to prevent such attacks, the specific attack would have to be known to keep it from entering a system, or it would be difficult to detect a novel attack [8]. An intelligent solution to IDS is necessary to predict novel attacks before the system becomes infected. Recent literature proposes using neural networks, such as ANN, DNN, CNN, and LSTM, for an IDS as a solution [9] [10].

978-1-7281-9615-2/20$31.00 ©2020 IEEE

An important component of IDS development is determining how to recognize certain messages to be anomalies, or, attacks [11]. Many recent IDS use neural networks, to train in predicting malicious behavior because of the capacity of the model to identify patterns [12] [13]. An important part of training a neural network is the pre-processing of training data [14]. There are many factors that can influence the outcome of training a neural network, which is why it is important to have a large training dataset. The trained network should also be validated on a separate dataset to ensure that it does not overfit the training data [15] [16].

Moreover, different scales of the attributes used to train networks can influence the outcome of the different features for prediction. Values on the scale of hundreds can have a larger influence on prediction than binary values. Normalization is an important pre-processing technique that changes all of the attributes to a similar scale so that prediction is more accurate. Attributes that are non-numeric also have to be altered to discrete values in the pre-processing phase as well.

Finally, there are many pre-processing algorithms used to determine the features and values for the features that are used in predicting within the IDS [17]. Some attributes are even removed entirely because they do not significantly impact the outcome of prediction, also known as feature selection [4]. Feature selection is an important tool that may increase an IDS's accuracy and create more effective systems [10] [18].

The current trends of neural networks in IDS motivate the analysis on the effects of data pre-processing. The discrepancy in performance of a pre-processing technique applied to various neural network architectures also serves as a driving factor.

A. Contributions

The major contributions of this paper are:

- Supplementing current research on the effects of various data pre-processing methods by creating a DNN that is trained and tested on the NSL-KDD network data.

- Measuring the importance of normalizing data for training a model to produce optimal accuracy, loss, F-Score, and AUC-ROC.
- Demonstrating the importance of pre-processing for neural network architectures.

II. Related Works

A. Datasets

Two of the commonly used benchmark datasets used for IDS are the NSL-KDD dataset and the CIDDS dataset. The details of each dataset are expanded upon in the following two subsections.

1) NSL-KDD: NSL-KDD is a dataset containing network information with both benign packets and anomalies. It is a modification of the KDD CUP '99 dataset and aims to solve the issues of its predecessor based on [19]. The NSL-KDD dataset improves on the original KDD CUP '99 dataset by removing redundant entries and dividing the dataset into separate training and testing sets [15]. Because of the sheer amount of records in the dataset, multiple experiments can be conducted on different sub-datasets, allowing a more holistic evaluation of IDS.

The significance of the dataset and its widespread use among researchers is due to the lack of availability of datasets for IDS experimentation. NSL-KDD fills the gap and enables testing with network data. Although it may not yet be fully representative of a real network, Tavallaee et al. [15] states it still acts as an effective benchmark for comparing IDS.

2) CIDDS: The CIDDS dataset consists of flow-based data and is generated using virtual networks and virtual machines. Malicious data is simulated by insider attacks and external server attacks. User behavior is simulated through Python scripts [20]. CIDDS is customizable and kept up-to-date [21]. The original dataset, CIDDS-001, captured 32 million data and features 92 attacks [22]. The dataset attempts to remove period repetitions and randomizes size and number of attachment files for emails while the malicious data includes port scans, brute force, and denial of service attacks [23].

The significance of CIDDS lies in its adaptability, complexity, and amount of available data. Researchers are able to customize the dataset and include more attacks to further test the capability of their IDS. By having more data at the user's disposal, they can generate a great deal of sub-datasets to test and demonstrate the functionality of their IDS.

3) Other Datasets: Some datasets that have also been used for testing IDS include the CIC-IDS 2017, CSE-CIC-IDS 2018, UNSW-NR15, CERT, and TWOS datasets [22] [24] [25]. Both the CIC-IDS 2017 dataset, which includes normal and attack traffic data collected over five days, and the CSE-CIC-IDS 2018 dataset, collected from Amazon's AWS LAN network, are real datasets collected by the Canadian Institute of Cybersecurity [22]. The UNSW-NR15 dataset was created using the IXIA Perfect Storm tool to emulate the environment for over 31 hours and is already split into training and testing sets [22] [25]. TWOS is a synthetic dataset that was obtained from a 5-day simulation where users sent normal or malicious packets, identified as masqueraders and traitors [26]. CERT is another dataset that contains synthetic data on USB device usage, websites visited, workstation logins, emails, and psychometric scores [27]. The dataset attempts to generate as close to a real dataset as possible using different graph, behavior, communication and other models for research [28]. The CIC-IDS 2017 and CSE-CIC-IDS 2018 datasets are only mentioned, but not used for evaluation, by Rashid et al. [22]. Aside from Toffalini et al. [29] using TWOS for IDS evaluation, CERT and TWOS were infrequently used in the papers reviewed. Finally, the UNSW-NR15 dataset is only used for evaluation by He et al. [25] but is used less frequently compared to CIDDS and NSL-KDD. Therefore, these datasets are not considered further in this paper.

B. Pre-Processing Techniques

IDS utilize machine learning to apply algorithms to analyze previous information and predict future attacks [17]. A focus on data pre-processing can greatly assist the improvement of proposed IDS models. Due to varying scales and strings values, a key step in pre-processing the dataset is transforming the strings into discrete values, and normalizing data [9]. Normalization and feature selection are valuable techniques for pre-processing data and can improve accuracy in benign and malicious packet detection [11].

1) Normalization: Obaid et al. [30] compares four types of normalization: min-max which preserves the relationship of the original values, ZScore which is useful for reducing the effect of outliers, decimal scaling, and log base 2 scaling.

Data must also be changed from discrete string values to discrete numeric values for processing. One approach is to map the different attributes to binary vectors that represent the different string values [31]. Other papers discuss ways of discretizing continuous-value features by dividing values into equal width groups [30] [32].

2) Feature Selection and Reduction: There are many algorithms that can be used to implement feature selection and reduction in an IDS. Some algorithms used for pre-processing data include: Naive Bayes, NBTree, J48 or C4.5 Decision Tree, Random Tree, Support Vector Machine (SVM), and BFGS quasi-Newton Backpropagation [17] [33]. Table **??** and Table **??** summarize techniques that papers use and their evaluated performance.

III. Our Approach

A. Pruning and Normalization Techniques

Results were collected using two validation datasets: KDD_Test+ and a pruned KDD_Test+. KDD_Test+ is an unaltered validation dataset provided by the NSL-KDD dataset that includes attacks not present in KDD_Train+ along with a slightly different distribution of attacks as a percentage of dataset entries. Thus the pruned KDD_Test+ set eliminates attacks not present in KDD_Train+. Testing against both validation sets allowed for analysis of normalization techniques against known attacks and novel attacks.

Normalization techniques, ZScore (Eq. 1) and Min-Max (Eq. 2), were applied to the training (KDD_Train+) and validation datasets (KDD_Test+ and pruned KDD_Test+) prior to passing them to the NN model. Key values such as mean, standard deviation, minimum, and maximum were taken per feature column from the training dataset and applied to the training and validation datasets. "None", as it appears in the results graphs indicates that no normalization techniques were applied.

$$ZScore: \quad \frac{InputData - Mean}{StandardDeviation} \tag{1}$$

$$Min - Max: \quad \frac{InputData - Min}{Max - Min} \tag{2}$$

B. Neural Network Architecture Selection and Training

We opt to use a DNN as the base architecture for our approach because of its nature as a simple model forming complex relationships. DNNs are subsets of ANNs that have multiple hidden layers and added complexity in connections they may form, making ANNs seem simplistic and less desirable in comparison [34] [35]. Further, there is a research gap in implementations of DNNs for IDS, which motivates the study of its potential in the field.

CNNs specialize in artificial vision and image processing, which deviate from the goal of IDS and data encountered [33] [36]. This mismatch in compatibility motivates against selecting a CNN. RNNs and LSTMs are cyclically connected and are more complex than other neural networks. Their accuracy for IDS implementations, however, does not greatly improve compared to the others [34] [37].

The DNN model tested consisted of an input layer of 27 features, two dense, hidden layers of 128 neurons each, and a 1 neuron output layer. The two hidden layers used a ReLu activation function and the output layer used a sigmoid activation function.

The Adam optimizer was used to adjust weights with a learning rate of 0.0001, $\frac{1}{10}$ of the default learning rate in TensorFlow. Reducing the default learning rate allowed for greater detail when plotting evaluated metrics per epoch, though at the cost of computation time.

C. Evaluated Metrics

Preprocessing techniques were evaluated based on the accuracy, binary cross-entropy loss, AUC-ROC, True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) for each epoch on the training dataset (KDD_Train+) and the validation datasets (KDD_Test+ and KDD_Test+ with attacks not present in KDD_Train+ removed) using functions built into Tensorflow. A callback function was used to compare both validation sets to the same trained model. F-Score was calculated later for each epoch using the following equation:

$$F - Score = \frac{2 * TP}{TP + TN + FP + FN}, \tag{3}$$

D. Testbed

Experimentation was conducted on a PC running Windows 10 and Python 3.7.4 with an AMD Ryzen 3900X CPU and an Nvidia GTX 1080 GPU. Pruning was and normalization techniques were applied using Pandas v1.0.3. The dataset was imported and the NN model built, trained, and validated using TensorFlow-GPU v2.1.0.

IV. Results

This section is organized by a discussion of each metric followed by a culminating analysis of the measured results. Each metric is compared between the two validation datasets and the three normalization techniques. Associated with each metric are graphs measuring the trends over all 200 epochs for each normalization technique and corresponding datasets. The quantitative results presented are the values at the 200th epoch in order to account for the fluctuation while the model is trained. Our analysis of the metrics highlights the differences in performance of the two validation sets. These results are summarized in the comprehensive table and analysis at the end of the section.

A. Accuracy

For accuracy, the different datasets tended to cluster together with the training having the highest accuracy all above 0.9, followed by the pruned validation dataset around 0.8 and finally the complete validation dataset around 0.7 for the final epoch values. As expected, the accuracy of the training datasets performed better than either of the two validation sets. The no normalization training set performed the best with 0.9466 accuracy, followed by ZScore with 0.9192 and then Min-Max with 0.9124. Between the two validation datasets, no normalization performed worse than the other normalization techniques with 0.7955 accuracy on the pruned dataset and 0.6936 on the complete dataset. For the pruned dataset, overall Min-Max normalization performed better with 0.8021 compared to 0.8009 with ZScore normalization. On the otherhand, ZScore performed better on the complete dataset with 0.7245 accuracy compared to 0.7074 with Min-Max normalization.

Figure 1 shows the accuracy trend over the 200 epochs the models were trained for using each of the three normalization techniques. The graph displays the training accuracy as well as the accuracy of the complete validation dataset and pruned validation dataset. As shown, each of the accuracies tend to follow the same general trend when evaluated on the same trained model with the same normalization technique. ZScore appears to level out and start converging the soonest, and has the smoothest curve when being trained, while Min-Max levels out and starts converging soon after ZScore does, but is less smooth with more jumps between epochs 1 and 25. Finally, no normalization converges the slowest and has the most bumpy graph for accuracy. In all three normalization techniques, the training accuracy performs the best then is followed by the pruned dataset by about 0.1 which is then followed by the complete dataset by about 0.1 again when the metric converges.

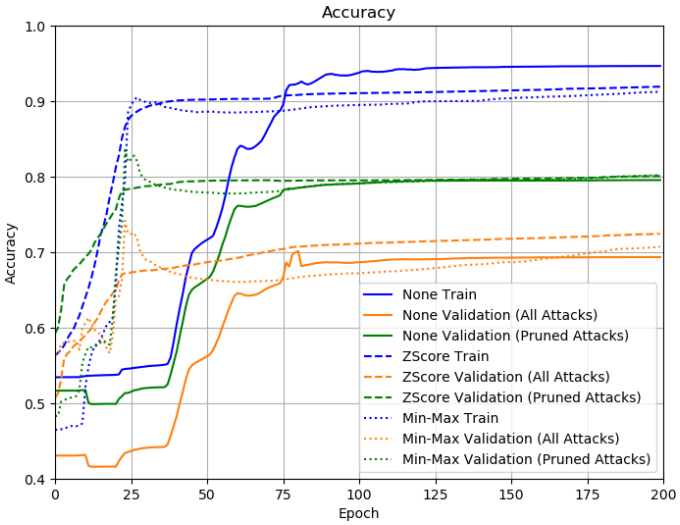


Fig. 1. Comparison of accuracy in training and validation datasets using different normalization techniques.

Based on these results, no normalization performs the best for the training dataset. This is because without any normalization, the model is trained on the raw data which allows for the most pattern recognition, but tends to overfit that data and perform worse on any new data. The accuracy metric is optimized for the pruned dataset, which contains no novel attacks, using Min-Max normalization and is optimized for the complete validation dataset, which does contain novel attacks, using ZScore normalization.

B. Binary Cross-Entropy Loss

Binary Cross-Entropy Loss, referred to as simply loss now on, had final values much closer to each other than accuracy. The best performing instance overall was also the no normalization training dataset with a final loss value of 0.5413. Also similar to the accuracy metric, the other two best performing instances were the other two training sets with a loss of 0.5561 for the ZScore training and a loss of 0.5793 for the Min-Max training. For the complete validation dataset, the loss metric was reduced the most to 0.5862 using ZScore normalization, followed by a loss of 0.5925 for Min-Max, and finally a loss of 0.5984 with no normalization. For the pruned validation dataset, the loss metric was reduced the most to 0.5943 with no normalization, followed by a loss of 0.5954 with ZScore, and finally a loss of 0.6023 with Min-Max.

Figure 2 shows the graph of the loss metric for each dataset and normalization technique over all 200 epochs the model was trained over. The graph of the loss metric for the datasets evaluated and trained with no normalization are the least smooth starting somewhat level for the first about 30 epochs then decreasing significantly for the following 40 epochs then starting to converge again. The loss graphs for the training and validation datasets that used ZScore and Min-Max are much more smooth throughout the entire training of the model, decreasing the entire time until they converge. Most of the loss metrics converge to a final value of about 0.6 with the training sets tending to be even lower.

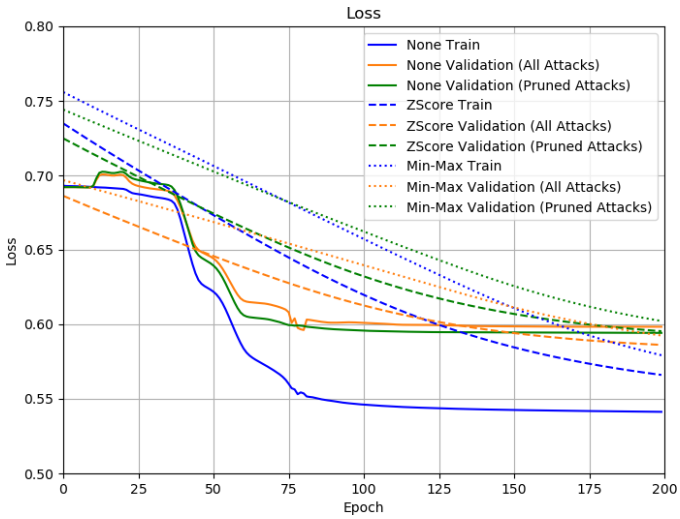


Fig. 2. Comparison of loss in training and validation datasets using different normalization techniques.

Although no normalization appears to have the best loss metric for the pruned dataset, the graph shows that it has converged while ZScore and Min-Max seem to still be decreasing a little for the pruned dataset. This means that 200 epochs might not have been enough for the two normalization techniques to converge. Similarly, although ZScore outperformed Min-Max normalization, the graphs show that the loss for Min-Max was still decreasing at a faster rate than ZScore. Because of this, it is unclear from the loss metric which normalization technique performs the best for our model and on the different datasets.

C. F-Score

The F-Score for the three normalization techniques will be examined. This metric provides a holistic overview of the model's performance as it varies with each technique. Figure 3 shows the F-Score for each normalization technique across 200 epochs. The trend for ZScore begins with a steep increase and then stabilizing from the 25th epoch onward. No normalization follows a similar pattern, but begins stabilizing at the 75th epoch. There is an exception when Min-Max normalization is used: there is a slight increase in F-Score until the 25th epoch, where the value drops for the 50 following epochs. The F-Score then slightly increases through the final epoch, meaning 200 epochs may not have been a sufficient amount of time to capture the convergence.

ZScore produced the highest F-Score, 0.7447, for the pruned validation dataset with Min-Max closely following at 0.7443. No normalization had the smallest F-Score value at 0.6283 This indicates no normalization is a better solution for cases where only known attacks will be encountered.

For the complete validation dataset, ZScore had the highest F-Score value, 0.6845. Following behind were Min-Max and

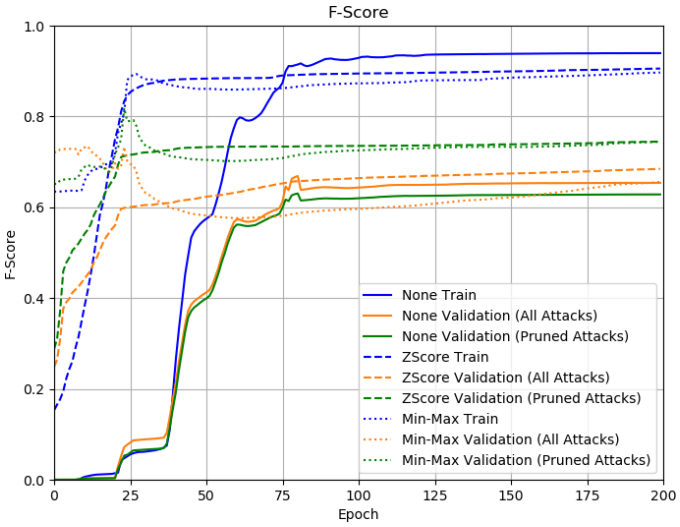


Fig. 3. Comparison of F-Score in training and validation datasets using different normalization techniques.

no normalization, with F-Scores of 0.6559 and 0.6538, respectively. The F-Score of ZScore normalization did not sharply drop as those of Min-Max and no normalization did. This demonstrates ZScore normalization builds a more adaptive model and will produce better results overall when detecting novel attacks.

D. Area Under the ROC Curve

With the exception of the two validation datasets with no normalization, all of the AUC-ROC values for the different models and datasets were above 0.9. As expected, the AUC-ROC values on the training datasets were higher than those of the validation datasets, with the highest being 0.9865 from the Min-Max normalized training dataset. The second highest AUC-ROC was 0.9731 from the ZScore normalized training dataset followed by the training dataset with no normalization at 0.9454. For the pruned validation dataset, the best AUC-ROC was 0.9393 from Min-Max normalization followed by ZScore with an AUC-ROC of 0.9333 and finally no normalization with an AUC-ROC of 0.8020. For the complete dataset, similar results were found with Min-Max performing the best at 0.9363, followed by ZScore with an AUC-ROC of 0.9328, and finally no normalization with 0.7544. The complete dataset performed worse than the pruned dataset by 0.0030 using Min-Max and by 0.0005 using ZScore which is smaller than the 0.0060 and 0.0035 difference between the normalization techniques for each validation dataset. Min-Max improved on no normalization by 0.1373 for the pruned dataset and by 0.1818 for the complete dataset while ZScore improved on no normalization by 0.1313 for the pruned dataset and 0.1784 for the complete dataset. Overall, Min-Max outperforms ZScore, but by a small amount, but both techniques are significantly better than using no normalization at all.

Figure 4 shows the graphs of the AUC-ROC for each normalization technique and dataset over all 200 epochs. For the majority of the epochs when comparing the same normalization technique, the training dataset performs the best for the majority of the epochs. For no normalization, all of the graphs start about the same for the first around 30 epochs then start to increase at different rates and level out. For ZScore, except for the very beginning where the validation datasets start higher, they have similar graphs with the validation datasets being lower than the training dataset. For Min-Max, the datasets start out about the same and increase at the same rate for the first about 20 epochs, then start to level out at different AUC-ROCs.

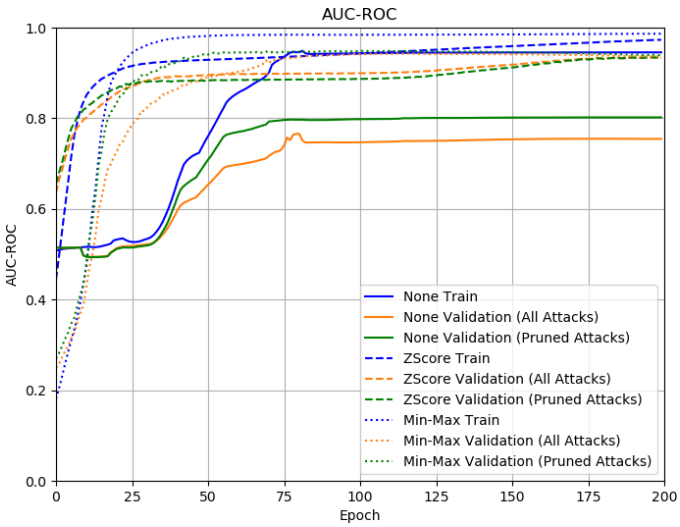


Fig. 4. Comparison of AUC-ROC in training and validation datasets using different normalization techniques.

These results demonstrate that the AUC-ROC using Min-Max and ZScore differs by less than 0.01, but Min Max performs slightly better than ZScore on both validation datasets. Also, the performance of AUC-ROC between the pruned and complete datasets differs by an even smaller value than the normalization techniques. Finally, using either Min-Max or ZScore shows an improvement on the AUC-ROC metric over using no normalization technique at all.

E. Summary of Results

Table I summarizes all of the metrics for each normalization technique and dataset of the 200th epoch. For each of the normalization techniques, the pruned validation set metrics generally outperformed the corresponding metrics of the complete validation sets. This can be attributed to the pruned validation set only containing attacks the model is already familiar with. The complete dataset includes attacks that the model was not trained with and would be harder for the IDS to detect as malicious which is why the pruned dataset performs better. In most cases, no normalization performed worse than both ZScore and Min-Max for each metric. Overall, ZScore and Min-Max perform better than no normalization in both datasets and the instances no normalization perform better

are due to the specific training of the model and not the normalization technique overall.

TABLE I
PERFORMANCE METRICS OF PRE-PROCESSING TECHNIQUES

Dataset	Normalization	Accuracy	Loss	F-Score	AUC
Pruned	None	0.7955	0.5943	0.6283	0.8020
	Min-Max	0.8021	0.6023	0.7443	0.9393
	ZScore	0.8009	0.5954	0.7447	0.9333
Complete	None	0.6936	0.5984	0.6538	0.7544
	Min-Max	0.7074	0.5925	0.6559	0.9363
	ZScore	0.7245	0.5862	0.6845	0.9328

Table II lists the TP, TN, FP, and FN for each dataset and normalization technique at the final epoch. The number of TN and FP remain the same between datasets for the corresponding normalization techniques because only the malicious instances were changed between the datasets.

TABLE II
PERFORMANCE METRICS OF PRE-PROCESSING TECHNIQUES

Dataset	Normalization	TP	TN	FP	FN
Pruned	None	5839	9111	600	3243
	Min-Max	5413	9660	51	3669
	ZScore	5458	9593	118	3624
Complete	None	6524	9111	600	6308
	Min-Max	6287	9660	51	6545
	ZScore	5458	9593	118	3624

Despite the shortcomings of the model, the effects of each pre-processing method can be identified and compared. Across all of the measured performance metrics for the pruned validation dataset, both Min-Max normalization ZScore normalization optimize each value. The complete validation dataset yields a different conclusion. Min-Max normalization marginally outperformed ZScore in Precision and AUC, while ZScore outperformed Min-Max normalization for the remaining metrics, but most especially in F-Score. This demonstrates using ZScore normalization will more correctly identify novel attacks compared to the other normalization methods. On a different DNN model better suited to IDS, pre-processing training data using ZScore normalization would enhance the performance greater than Min-Max or no normalization would. The ZScore normalization also demonstrates a smoother, more stable curve over the epochs compared to both Min-Max and no normalization. This is indicative of consistency in growth rate and convergence of each metric when implementing ZScore normalization. Implementing Min-Max normalization for a DNN-based IDS would produce results similar to ZScore, since the metric values were marginal. The percentage improvement from implementing ZScore or Min-Max over no normalization is calculated for accuracy, F-Score, and AUC-ROC. For the pruned dataset, using ZScore normalization induced a 0.68% increase in accuracy, a 0.19% increase in loss, a 18.53% increase in F-Score, and a 16.37% increase in AUC-ROC when compared to having no normalization technique in place. Min-Max normalization produced a 0.83% increase for accuracy, a 1.35% increase in loss, a 18.46% increase in F-Score, and a 17.12% increase in AUC-ROC compared to no normalization. For the complete dataset, using ZScore normalization saw a 4.46% increase in accuracy, a 2.04% decrease in loss, a 4.70% increase in F-Score, and a 23.64% increase in AUC-ROC. Min-Max normalization produced a 1.99% increase in accuracy, a 1.00% decrease in loss, a 0.32% increase in F-Score, and a 23.65% increase in AUC-ROC. It can be concluded then, that using either of these normalization techniques increases performance compared to no normalization.

V. CONCLUSION

Normalization and other pre-processing techniques applied to the data used for training an IDS are important for optimizing the performance metrics. We propose a DNN using 27 input features for binary classification trained using the NSL-KDD dataset. As expected, the experimentation on the pruned dataset outperforms that of the complete dataset across most metrics. Our proposed model determines for the complete dataset that ZScore normalization, and Min-Max normalization to a lesser degree, improve the performance of the proposed IDS compared to no normalization. Our results demonstrate that ZScore improves on no normalization by 4.46% for accuracy, 2.04% for loss, 4.70% for F-Score, and 23.64% for AUC-ROC. Min-Max normalization presents similar improvements with an increase in accuracy at 1.99%, decrease in loss at 1.00%, increase in F-Score at 0.32%, and increase in AUC-ROC at 23.65%. Implementing ZScore normalization as a pre-processing step can improve the performance of DNN-based IDS.

REFERENCES

[1] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41–50, 2018.

[2] A. Kim, M. Park, and D. H. Lee, "Ai-ids: Application of deep learning to real-time web intrusion detection," *IEEE Access*, vol. 8, pp. 70 245–70 261, 2020.

[3] E. Benkhelifa, T. Welsh, and W. Hamouda, "A critical review of practices and challenges in intrusion detection systems for iot: Toward universal and resilient systems," *IEEE Communications Surveys Tutorials*, vol. 20, no. 4, pp. 3496–3509, 2018.

[4] H. BENADDI, K. IBRAHIMI, and A. BENSLIMANE, "Improving the intrusion detection system for nsl-kdd dataset based on pca-fuzzy clustering-knn," in *2018 6th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, 2018, pp. 1–6.

[5] C. Ma, X. Du, and L. Cao, "Analysis of multi-types of flow features based on hybrid neural network for improving network anomaly detection," *IEEE Access*, vol. 7, pp. 148 363–148 380, 2019.

[6] O. Y. Al-Jarrah, O. Alhussein, P. D. Yoo, S. Muhaidat, K. Taha, and K. Kim, "Data randomization and cluster-based partitioning for botnet intrusion detection," *IEEE Transactions on Cybernetics*, vol. 46, no. 8, pp. 1796–1806, 2016.

[7] M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan, "Building an intrusion detection system using a filter-based feature selection algorithm," *IEEE Transactions on Computers*, vol. 65, no. 10, pp. 2986–2998, 2016.

[8] A. Nisioti, A. Mylonas, P. D. Yoo, and V. Katos, "From intrusion detection to attacker attribution: A comprehensive survey of unsupervised methods," *IEEE Communications Surveys Tutorials*, vol. 20, no. 4, pp. 3369–3388, 2018.

[9] B. Ingre and A. Yadav, "Performance analysis of nsl-kdd dataset using ann," in *2015 International Conference on Signal Processing and Communication Engineering Systems*, 2015, pp. 92–96.

[10] A. Gül and E. Adalı, "A feature selection algorithm for ids," in *2017 International Conference on Computer Science and Engineering (UBMK)*, 2017, pp. 816–820.

[11] L. Hakim, R. Fatma, and Novriandi, "Influence analysis of feature selection to network intrusion detection system performance using nsl-kdd dataset," in *2019 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*, 2019, pp. 217–220.

[12] L. N. Tidjon, M. Frappier, and A. Mammar, "Intrusion detection systems: A cross-domain overview," *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 3639–3681, 2019.

[13] S. Tayeb, N. Raste, M. Pirouz, and S. Latifi, "A cognitive framework to secure smart cities," in *MATEC Web of Conferences*, vol. 208, 2018, p. 1.

[14] S. Tayeb, M. Pirouz, B. Cozzens, R. Huang, M. Jay, K. Khembunjong, S. Paliskara, F. Zhan, M. Zhang, J. Zhan *et al.*, "Toward data quality analytics in signature verification using a convolutional neural network," in *2017 IEEE international conference on big data (Big Data)*. IEEE, 2017, pp. 2644–2651.

[15] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the kdd cup 99 data set," in *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, 2009, pp. 1–6.

[16] S. Tayeb, M. Pirouz, J. Sun, K. Hall, A. Chang, J. Li, C. Song, A. Chauhan, M. Ferra, T. Sager *et al.*, "Toward predicting medical conditions using k-nearest neighbors," in *2017 IEEE International Conference on Big Data (Big Data)*. IEEE, 2017, pp. 3897–3903.

[17] R. Thomas and D. Pavithran, "A survey of intrusion detection models based on nsl-kdd data set," in *2018 Fifth HCT Information Technology Trends (ITT)*, 2018, pp. 286–291.

[18] S. Sarvari, N. F. Mohd Sani, Z. Mohd Hanapi, and M. T. Abdullah, "An efficient anomaly intrusion detection method with feature selection and evolutionary neural network," *IEEE Access*, vol. 8, pp. 70 651–70 663, 2020.

[19] J. Mchugh, "Testing intrusion detection systems: a critique of the 1998 and 1999 darpa intrusion detection system evaluations as performed by lincoln laboratory," *ACM Transactions on Information and System Security (TISSEC)*, vol. 3, no. 4, pp. 262–294, 2000.

[20] M. Ring, S. Wunderlich, D. Grüdl, D. Landes, and A. Hotho, "Flow-based benchmark data sets for intrusion detection," in *Proceedings of the 16th European Conference on Cyber Warfare and Security (ECCWS)*. ACPI, 2017, pp. 361–369.

[21] S. A. Althubiti, E. M. Jones, and K. Roy, "Lstm for anomaly-based network intrusion detection," in *2018 28th International Telecommunication Networks and Applications Conference (ITNAC)*, 2018, pp. 1–3.

[22] A. Rashid, M. J. Siddique, and S. M. Ahmed, "Machine and deep learning based comparative analysis using hybrid approaches for intrusion detection system," in *2020 3rd International Conference on Advancements in Computational Sciences (ICACS)*, 2020, pp. 1–9.

[23] M. Ring, S. Wunderlich, D. Grüdl, D. Landes, and A. Hotho, "Creation of flow-based data sets for intrusion detection," *Journal of Information Warfare*, vol. 16, pp. 40–53, 2017.

[24] A. Verma and V. Ranga, "On evaluation of network intrusion detection systems: Statistical analysis of cidds-001 dataset using machine learning techniques." *Pertanika Journal of Science & Technology*, vol. 26, no. 3, 2018.

[25] W. He, H. Li, and J. Li, "Ensemble feature selection for improving intrusion detection classification accuracy," in *Proceedings of the 2019 International Conference on Artificial Intelligence and Computer Science*, ser. AICS 2019. New York, NY, USA: Association for Computing Machinery, 2019, p. 28–33. [Online]. Available: https://doi.org/10.1145/3349341.3349364

[26] A. Harilal, F. Toffalini, J. Castellanos, J. Guarnizo, I. Homoliak, and M. Ochoa, "Twos: A dataset of malicious insider threat behavior based on a gamified competition," in *Proceedings of the 2017 International Workshop on Managing Insider Security Threats*, ser. MIST '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 45–56. [Online]. Available: https://doi.org/10.1145/3139923.3139929

[27] S. C. Roberts, J. T. Holodnak, T. Nguyen, S. Yuditskaya, M. Milosavljevic, and W. W. Streilein, "A model-based approach to predicting the performance of insider threat detection systems," in *2016 IEEE Security and Privacy Workshops (SPW)*, 2016, pp. 314–323.

[28] J. Glasser and B. Lindauer, "Bridging the gap: A pragmatic approach to generating insider threat data," in *2013 IEEE Security and Privacy Workshops*, 2013, pp. 98–104.

[29] F. Toffalini, I. Homoliak, A. Harilal, A. Binder, and M. Ochoa, "Detection of masqueraders based on graph partitioning of file system access events," in *2018 IEEE Security and Privacy Workshops (SPW)*, 2018, pp. 217–227.

[30] H. S. Obaid, S. A. Dheyab, and S. S. Sabry, "The impact of data pre-processing techniques and dimensionality reduction on the accuracy of machine learning," in *2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON)*, 2019, pp. 279–283.

[31] K. Wu, Z. Chen, and W. Li, "A novel intrusion detection model for a massive network using convolutional neural networks," *IEEE Access*, vol. 6, pp. 50 850–50 859, 2018.

[32] D. H. Deshmukh, T. Ghorpade, and P. Padiya, "Improving classification using preprocessing and machine learning algorithms on nsl-kdd dataset," in *2015 International Conference on Communication, Information Computing Technology (ICCICT)*, 2015, pp. 1–6.

[33] B. Subba, S. Biswas, and S. Karmakar, "A neural network based system for intrusion detection and attack classification," in *2016 Twenty Second National Conference on Communication (NCC)*, 2016, pp. 1–6.

[34] R. K. Vigneswaran, R. Vinayakumar, K. P. Soman, and P. Poornachandran, "Evaluating shallow and deep neural networks for network intrusion detection systems in cyber security," in *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2018, pp. 1–6.

[35] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, pp. 41 525–41 550, 2019.

[36] R. Blanco, P. Malagón, J. J. Cilla, and J. M. Moya, "Multiclass network attack classifier using cnn tuned with genetic algorithms," in *2018 28th International Symposium on Power and Timing Modeling, Optimization and Simulation (PATMOS)*, 2018, pp. 177–182.

[37] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, pp. 21 954–21 961, 2017.

THE USE OF AGILE METHODOLOGY FOR PORTING ANALOG AND MIXED-SIGNAL CIRCUITS BETWEEN DIFFERENT TECHNOLOGY NODES

Ebru Dalbudak
Computer Science and Engineering, Université du Québec en Outaouais
Gatineau, QC, J8Y 3G5, Canada
dale02@uqo.ca

Karim Baratli
Computer Science and Engineering, Université du Québec en Outaouais
Gatineau, QC, J8Y 3G5, Canada
bark07@uqo.ca

Yousef Fouzar
Computer Science and Engineering, Université du Québec en Outaouais
Gatineau, QC, J8Y 3G5, Canada
yfouzar@yahoo.ca

Bachir Lakhssassi
Computer Science and Engineering, Université du Québec en Outaouais
Gatineau, QC, J8Y 3G5, Canada
Lakb01@uqo.ca

Karim El Guemhioui
Computer Science and Engineering, Université du Québec en Outaouais
Gatineau, QC, J8Y 3G5, Canada
Karim.elguemhioui@uqo.ca

Ahmed Lakhssassi
Computer Science and Engineering, Université du Québec en Outaouais
Gatineau, QC, J8Y 3G5, Canada
Ahmed.lakhssassi@uqo.ca

Abstract—With the growing need for increased capacity, speed and capabilities, the complexity of Integrated Circuit (IC) designs continue to increase. With the increased level of complexity and market pressures to deliver better results faster and cheaper, IC companies are forced to innovate and introduce efficiencies into their development process. To improve IC design productivity and time-to-market, this paper introduces an innovative Electronic Design Automation (EDA) tool that will allow the replication of an existing layout in different technology nodes by automatically porting analog and mixed-signal circuits. Furthermore, to better manage the complexity of IC development projects, the paper suggests the use of iterative and incremental approach of agile methods by presenting a case study on an EDA tool development project.

Keywords— IP Porting, IC design productivity, EDA, Technology Migration, Layout Migration

I. INTRODUCTION

Since the invention of the first Integrated Circuit (IC) in 1958, the number of transistors per IC has doubled approximately every two years for decades as per Intel's Gordon Moore's observation, referred to as "Moore's Law" [1]. Each generation of semiconductor process improvements shrinks the size of the transistors and wires and increases the number of transistors and wires that can fit on a chip dramatically. Consequently, this allows the chips to run at higher frequencies and then the new designs get more sophisticated.

With the increased level of complexity in designs, comes greater design challenges and a growing risk of failing to ensure time to market, for semiconductor companies [2].

As the process technology advances, IC design rules[1] that are provided by semiconductor manufacturers get more complicated and usually they are not compatible with previous generations [3]. According to Zhu et al.[4], "manufacturing processes are updated every 18 months each with a different set of design rules", and companies need to offer different versions for different foundries. The standard cell library, which consists of standardized layout specific components, for each process update also increases [5]. As a result, regenerating a new standard cell library for each process update gets more costly and time consuming [5, 6]. Moreover, the circuits need to migrated to work with the new target technology and this requires a tedious, often manual process of iterating layout changes, extraction and simulation until desired characteristics are met [1].

To maximize their design investments and reduce their design duration, Intellectual Property (IP) providers reuse blocks and hard IP whenever possible. At the beginning of or during a new IC design project, it is often necessary to use and re-use available IP and circuits from previous projects, instead of designing everything from scratch [2, 7]. Design reuse makes it faster and cheaper to design and the time lost by completely re-designing an existing circuit in another technology process can then be spent on new parts of the system that add extra functionality and/or increase performance and productivity [8].

Furthermore, since IC companies choose to outsource the fabrication of their semiconductor chips to foundries (by going fabless), they need to migrate their existing circuit designs "between process technologies and within internal and external foundry processes" because, frequently, foundries

[1] Design rules determine whether the physical layout of a given circuit satisfies a series of recommended parameters to ensure that most of the parts will work correctly.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

differ in their versions even if they use same process technology node [7]. Particularly, the number of metal layers, material table, etc.

Hence, there is a growing need in the semiconductor market to migrate an existing circuit layout between different technology nodes and foundries in a reliable, fast and cost efficient manner [3, 4, 8, 9]. To produce designs quickly and to address design challenges, IC companies have been increasingly making use of the EDA tools [1].

This paper introduces two innovative approaches that will improve IC design productivity and hence will facilitate faster semiconductor development and time-to-market. First, it introduces a new EDA tool in microelectronics which aims to automatically port analog and mixed-signal circuits between technologies nodes in the hard IP market. Next, it demonstrates the use of agile methodology in the EDA tool development project to promote the adaptive planning and evolutionary development and delivery and presents how a module of this tool was developed in short and iterative development cycles.

The article begins with introducing concepts relevant to the study as an overview. Next, it provides a detailed discussion on layout porting and its challenges. The paper continues with the introduction of the new EDA tool called OPIC that can help better manage the complexity in microelectronics for IP design porting of analog and mixed signal circuits between technology nodes. It proceeds with the case study on the use of agile methods during the development of a Design Rule Check (DRC) clean-up module of the OPIC tool. And it ends with concluding remarks.

II. Background / Related Work

There are studies that were focused on IP reuse and technology migrations due to their expected benefits.

Sobe et al. [10] presents a method, that reflects common practice of analog design divided into design data conversion and sizing by optimization, to convert and optimize a circuit topology. Fu et al. [11] introduce a new rectangular topological layout for topology-driven cell migration. Kar et al. [12] talks about the technology migration process and introduce a new layout preserving migration tool called: TECHMIG. Francken et al. [2] presents a methodology for the migration of analog circuit design from one technology process to another, considering both the sizing and the layout phase. Diodato et al. [13] discuss porting of 2.5 micron CMOS VLSI design into 1.75 micron design rules. They were able to implement the new design in 33% less time and achieved smaller size in design rules by 19% and higher operating frequency by 43%. Dornelas, Helga et al. [6] introduced a new technology migration methodology for analog IC design and instead of doing the migration manually or at netlist level, they performed automated migration on a schematic level, followed by a robustness verification with the usage of MunEDA tools [6, 10].

However, literature review has revealed that even though there is abundant research in the field of IP reuse, design migration and layout porting, their use in analog and mixed-signal circuits is not as common [3].

The literature review on the EDA tools and the use of agile in IC development projects has proven that the use of agile methods in hardware development projects is not as common as in software development projects [14]. Innovation is a great challenge in the EDA industry and to provide faster, better, and cheaper products, the major players in the industry focus on improving existing tools and methodologies [15]. According to Bryant et al. “The quality of the design tools and associated methodologies determine the design time, performance, cost and correctness of the final system product” [16].

III. Layout Porting

As Francken et al. stated “porting successful designs from one technology to another can distinctly reduce the design cycle” [2]. The aim of layout porting between technological process nodes is to be able replicate an existing layout (reference design) in different technology nodes by modifying or adapting it to a newer process design rules by taking advantage of the reusability of the layout resources accumulated from the initial process. Semiconductor foundries continue to innovate by continuous scaling of transistors. There is a growing need to redesign mixed signal circuits for new technological processes. The analog domain still lacks automation, which resulted in a manual design of analog circuits each time to migrate from one technology to another. While doing that the ultimate goal is to “achieve high quality layout with high integration and low power consumption and to shorten the design-time-to-market” [8]. This way, instead of spending time and effort from scratch to completely redesign an entire circuit, semiconductor companies can focus on innovation by adding new functionalities into their designs and focusing on increasing performance [2].

IC products are usually mixed signal designs, in other words, they comprise of a combination of digital and analog functional blocks and circuits and the analog portions of these ICs are the most difficult portions of the chip to migrate between foundries and technology nodes [3, 17, 18]. In digital designs, reuse of IPs has already been well established thanks to fully automated digital workflow [8]. However, in analog circuit designs, reuse of IP have challenges, such as “coping with trade-offs among analog specific requirements, such as noise, linearity, gain, supply voltage, speed, power consumption, self-heating, etc” [8].

Porting an existing design from technology A to technology B usually follows a two step design flow. In the first step, the given schematics and topologies are converted from the source to target technology and existing layout and mapping layers

are migrated by recognizing different structure and device placements; preserving all floorplan, matching, routing and interconnect within the given layout [7, 19]. In the second step, these circuits and IPs are sized for the new target specifications and optimized for the new target process technology. Then, Design Rule Check (DRC) verification is conducted to preserve all matching within given layout and all routing and interconnect and Layout Versus Schematic (LVS) will be performed to ensure that the layout conforms to the rules required for faultless fabrication and performance [4].

At the present moment, in the IC design process, schematics and layout design phases are carried out sequentially. We start with the original reference design (for example 65nm), when the technology changes and we want for example 45nm technology node, we need to first migrate the specs (voltage is different, Vref(min, max)); get the schematics and circuit topologies for the new technology; translate schematics and topologies from 65nm process technology to 45nm process technology (via script); port the schematics from 65nm to 45nm; verify schematics; then use the schematics to generate/port the layout from 65nm to 45nm and verify the layout created in 45nm. As can be seen, in this type of a design, we do not take advantage of IP-reuse; design that works in one technology will not work in the other technology. So, IP has to be completely redesigned when the technology changes, which leads to more errors in the new design, added time, cost and human resources. Just to give an idea, today, about 100 engineers work together to design a new IP.

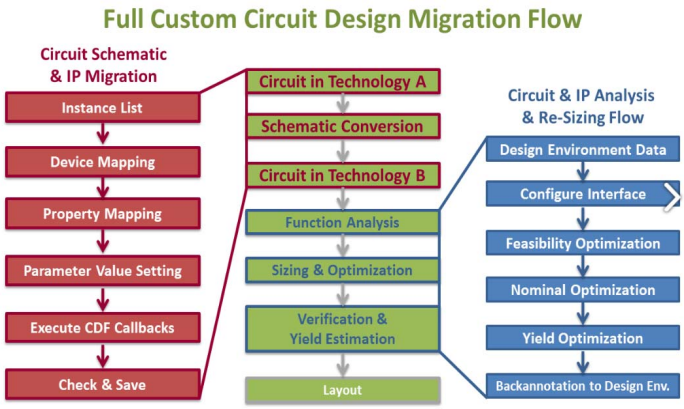


Figure 1: 2 steps of an IP Migration & Porting Flow [4]

IV. New Layout Porting Tool: OPIC

This paper introduces a new automated layout porting tool to increase the effectiveness and efficiencies of the design conversion process by reducing the human resources required and by shortening the porting time while reducing the failures during conversion.

There are some target layout recognition and conversion tools available commercially. However, these tools perform porting only when the reference design and the target technology is highly specialized (e.g. microcontrollers, etc.) and homogeneous (e.g. FPGA to FPGA).

Available commercial tools do not include features like automatic porting between technology nodes and foundries for analog and mixed-signal circuits; they do not re-use the existing layout by making changes on it; instead they regenerate the layout from scratch using traditional manual porting methods. Given the size and complexity of each migration, traditional manual porting methods are too time-consuming and costly. They are also risky and human intervention could impact the quality of the IP resulting in the loss of work that was already done in previous technology. IC designers need to reduce a risk of having costly design bug during the manual porting process from scratch. Therefore, the new layout needs to be retested rigorously for reliability, which impacts the IP delivery timelines. Moreover, instead of losing their precious time on repetitive manual tasks, IC designers should rather focus on using their creativity and precision crafting their designs.

A tool that would automatically migrate any IP design and porting flow with minimum human intervention from one foundry specific technology to another would be invaluable for any IP provider. By using this kind of an automated tool, semiconductor companies can increase the efficiency of their IC design process by reducing the porting time and the risk of creating bugs or errors, as well as the costs.

OPIC, which stands for "Outil de Portabilite entre technologies des Circuits Integrés/ Porting Tool for Integrated Circuits", is a new EDA tool that is designed for hard IP market to automatically perform the porting tasks with minimum human intervention and replicate an existing analog/mixed-signal circuit design (reference design) within different foundries and technology nodes.

The OPIC tool responds as well to an immediate need in the field of semiconductors and historical transition from traditional planar CMOS transistors to FinFETs, which is pending any adequate solution. The tool will be particularly useful in a variety of different important sectors, including IP providers, silicon foundries etc. within the semiconductor industry.

This automated porting tool is a type of back-end design tool, which aims to provide a quick and robust analysis of IP using existing layout, floorplan and routing between the blocks and will generate a DRC (Design Rule Check) and LVS (Layout Versus Schematic) clean layout in the targeted technology with minimum human intervention. The tool will port the graphical patterns (polygons) for each transistor and wire on the IC from one technology to another while preserving the key characteristics of an existing layout including matching of critical components and their relative placement in order to create a layout design that will function properly when fabricated.

The tool follows two major steps; first it will create a DRC and LVS database in technology B and then will perform circuit optimization and reflect schematic changes onto a clean layout.

The OPIC tool can be used to do the circuit and IP porting of a layout in source technology A to target technology B by converting the given schematics and topologies from the source to the target technology. It checks the geometries against geometry design rules, detects design rule violations according to specified IC parameters and automatically cleans them iteratively. As a final product, it generates an IC design in GDSII format, which is ready to be shipped to foundry for manufacturing.

The tool takes an existing design layout in a given technology (technology A), which consists of network of transistors and described by a geometrically exact Graphics Data System II (GDSII) file, which is the de facto industry standard for data exchange of IC design layout, and generates a new layout design in a different technology node (technology B) by using the design spec for the new layout in technology B and the technical mappings (for both A and B). which can be reviewed by a physical layout designer before it gets sent out to the foundry for manufacturing.

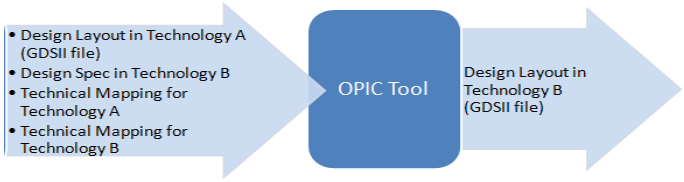


Figure 2: OPIC tool

To perform porting between two technological process nodes (from A to B), we need to have the following as input and follow the following steps listed below.

Input: Existing Layout Design in technology A in GDSII format, Design spec for the new layout in technology B, Technical mapping for technology A, Technical Mapping for technology B

Steps:

1. Open the given layout design in technology A with spec B and Existing Layout Design in technology A in GDSII format, Design spec for the new layout in technology B, Technical mapping for technology A, Technical Mapping for technology B migrate the layout into new layout design in technology B
2. Run the EDA tool called Design Rule Check (DRC) on Layout of technology B to identify the violations of design rules
3. Correct all the violations automatically and iteratively until the violations are cleaned
4. Obtain 100% accurate (target) layout for new design in technology B

Output: New layout design in technology B in GDSII format (binary file), which is ready to be given to IC foundries

We can run simulations on this new layout design in technology B to test it further and send this layout to a Physical layout designer for verification before the design gets sent to the foundry.

V. Case Study: The Use Of Agile Methodology In OPIC Project

OPIC tool is a long term project developed in phases, module by module. DRC rule cleanup is the first module of OPIC that was implemented successfully tested. DRC checks if the layout satisfies a set of design rules required for manufacturing. If the design rules are violated, the chip may not be functional.

OPIC DRC cleanup tool was implemented using JAVA and it is compatible with major EDA tools, such as Cadence, Mentor Graphics and Synopsys. The tool has been tested to ensure that it can correct DRC errors on a given layout (in GDSII file format) and the results have shown that the tool could reduce the tedious work of DRC checking from 25 hours of manual porting labor to 30 sec as demonstrated in the table below.

Reference Design	# and type of Errors	Manual correction	OPIC DRC Clean-up Tool
BIAS_GENERATOR	56 (7 type violations: Overlap, spacing, width, extension, clearance, etc.)	50 minutes	30 seconds (24 CPU)
Oscillator (7x7 inverters, 0,18nm)	378 (7 type violations: Overlap, spacing, width, extension, clearance, etc.)	240 minutes (no MACRO used, assuming different errors and no compaction problem)	183 seconds (24 CPU)

65nm and 28nm technology nodes were the two platforms used for verification. A reference design including various analog and mixed-signal circuits was built in 65nm technology and OPIC DRC tool was used to determine if the layout satisfied a set of rules required for manufacturing to migrate it to 28nm.

OPIC DRC clean-up is based on a proprietary algorithm (similar to the backtracking/shrinkage algorithm) and it includes 10 categories of common DRC errors that are related to spacing between metals, minimum width, etc. These errors are PO.C.2 (Polysilicon Oxide layer, Clearance), M1.W.1 (Metal Layer – Width), NW.W.1 (N-Well Layer, Width), NW.W.2 (N-Well Layer, Width), NW.S.1 (N-Well Layer, Spacing), NW.S.2 (N-Well Layer, Spacing), M1.S.1 (Metal Layer, Spacing), CO.E.2 (Contact Enclosure), NP.E.1 (N-doped Poly/substrate, Enclosure), PO.O.1 (Polysilicon Oxide layer, Overlap). An engine for correcting each error category was created. The tool recognizes these 10 different DRC errors and based on the error category it selects the engine that will correct the error. Until all the errors are cleaned-up and the DRC is clean, the tool runs iteratively, without human intervention.

OPIC DRC tool was developed by using a combination of agile software development and agile project management methodologies. Instead of defining the whole requirements, resources, technologies and tools upfront by using a plan-driven methodology like a waterfall, an agile development process was adopted, where the requirements, design and the process emerge in parallel during short, iterative cycles called *sprints*.

As shown in Figure 3 [20], the software development was based on an iterative life cycle and in one iteration cycle, an increment of the software was analyzed, designed, built, tested and integrated on a continuous basis until the full solution was realized [21]. At the end of each iteration, a subset of functionality was delivered to stakeholders for review and the software was built incrementally, piece by piece [22]. This way, the overall risk was minimized and the project could

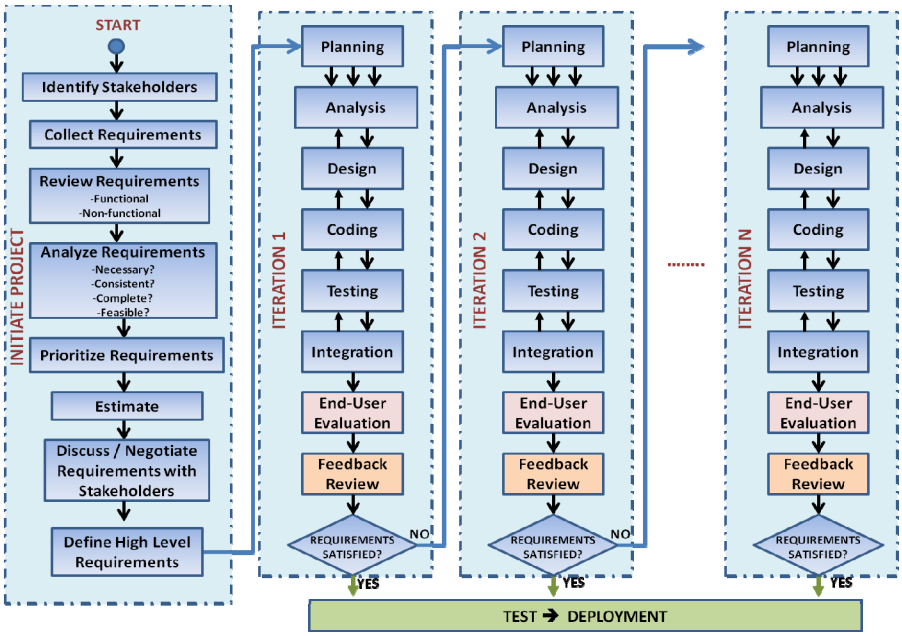


Figure 3: Agile Development methodology used in the OPIC project

adapt to changes quickly.

To be more specific, a hybrid agile approach of eXtreme Programming (XP) and SCRUM (where XP engineering practices are implemented in SCRUM sprints) were chosen and adopted in this project. According to Agile Alliance[2], "XP is an agile software development framework that aims to produce higher quality software, and higher quality of life for the development team" while SCRUM is a project management approach.

XP focuses on the developer side of the work with a set of engineering practices it mandates such as test driven development, pair programming, continuous integration, etc. which are very essential for building good quality software. SCRUM focuses on the project management side on what needs to be done. Since SCRUM does not have any engineering practices and XP does not have any management practices, XP and Scrum are well aligned and complement each other.

In the OPIC DRC tool project, the best practices recommended by XP were used for agile software development, such as simple design, pair programming, and small, frequent releases. SCRUM project management framework was used to manage the steps taken to develop software in conjunction with the use of XP to ensure the quality of the software.

This project was led by a project manager, who was responsible for coordinating the work of the cross-functional, self-organizing Scrum team of 5 people and for ensuring that tasks were completed within the scheduled time. The work was broken down to fit 3-week long sprints. Each sprint includes the traditional phases of software development: requirements analysis, design, coding, testing.

As shown in Figure 4, at the beginning of each sprint, a Sprint Planning meeting was held to plan the work to be done during the given sprint. In this meeting, a Sprint backlog that contains all the user stories (work items) to be implemented during the sprint was created, reviewed and prioritized. The team had daily scrum meetings of maximum 15 minutes to review the progress and productivity, to synchronize the information and to address the blocking issues. Team members worked on the user stories until completed. At the end of each sprint, a Sprint review meeting was held to demonstrate the work that was completed during the sprint and to gather feedback. The Sprint retrospective meeting was held shortly after that to reflect

[2]A non-profit organization dedicated to promoting agile methodologies at agilealliance.org

upon the last sprint in order to identify what went well and possibilities for improvement.

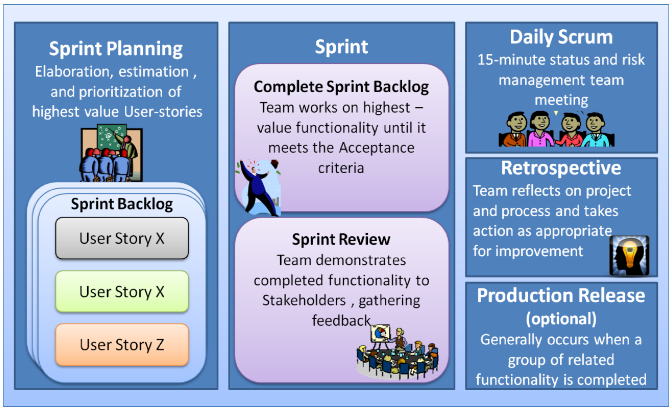


Figure 4: Scrum model that was followed in the OPIC project

Agile methodology was successfully adopted in this project to deliver the OPIC DRC tool in an iterative, incremental approach. Combining two agile approaches XP (for development practices) with Scrum (for project management) proved to be effective. Planning and big-picture view was crucial. Everyone in the team needed to understand what they were working towards for each piece to be able to come together in the end. Establishing an upfront, common understanding of the criteria for done-ness for each user story, for each sprint and for the release was very necessary. Clear and open communication was paramount. Note that agile mindset had to be understood and embraced by everyone in the team it took some time to refine the Scrum model to the point where team members were comfortable and productive.

VI. Conclusion

This paper has introduced a new EDA tool called OPIC that will automate porting of analog and mixed signal circuits from one technology node to another. Due to the complexity of recent advanced technology nodes, no automatic or pseudo-automatic tool has been developed for IP porting yet. The paper also presented a real-life example of adopting a hybrid agile methodology (combination of XP and SCRUM) in an IC development project by using the OPIC DRC tool project as a case study.

Automated OPIC DRC cleanup tool represents a major step for layout porting. This tool could be used as a standalone product and integrated with major commercially available tools to perform a DRC cleanup. It will allow designers save a considerable amount of time and use this time to focus on precision crafting their designs without sacrificing creativity to repetitive manual tasks.

Once fully implemented, OPIC tool will perform both the DRC and LVS corrections, automatically. So, the next step is to create an LVS tool, which takes a schematic to identify and check IC electrical connectivity against IC schematics and makes corresponding changes on the layout.

The need for increased capacity for more complex chips, speedups and new capabilities is growing and the time-to-market is getting shorter, increasing design productivity by utilizing IP-reuse and EDA tools will continue to play a critical role in semiconductor design and development.

References

1. Birnbaum, M.D., *Essential Electronic Design Automation (EDA)*. Prentice Hall PTR/Pearson Education, 2004.
2. Francken, K. and G. Gielen. *Methodology for analog technology porting including performance tuning*. in *Circuits and Systems, 1999. ISCAS'99. Proceedings of the 1999 IEEE International Symposium on*. 1999. IEEE.
3. Scheible, J. and J. Lienig. *Automation of analog IC layout: challenges and solutions*. in *Proceedings of the 2015 Symposium on International Symposium on Physical Design*. 2015. ACM.
4. Zhu, J., F. Fang, and Q. Tang, *Calligrapher: a new layout-migration engine for hard intellectual property libraries.* IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2005. **24**(9): p. 1347-1361.
5. ME, M.N.R.K., *DEVELOPMENT OF HIGH PERFORMANCE STANDARD CELL LIBRARY IN UMC180nm TECHNOLOGY.* International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE), 2014. **3**(9).
6. Dornelas, H., A. Schmidt, G. Strube, and E. Fabris, *New Technology Migration Methodology for Analog IC Design.*
7. MunEDA *Automated Design Migration & IP Porting Flow for Custom Circuits*. 2014. https://www.muneda.com/Applications_Custom-Circuit-Design-Migration-&-IP-Porting.
8. Nojima, T., S. Nakatake, T. Fujimura, K. Okazaki, Y. Kajitani, and N. Ono. *Adaptive porting of analog IPs with reusable conservative properties*. in *IEEE Computer Society Annual Symposium on Emerging VLSI Technologies and Architectures (ISVLSI'06)*. 2006. IEEE.
9. Qian, L., Z. Bi, D. Zhou, and X. Zeng, *Automated technology migration methodology for mixed-signal circuit based on multistart optimization framework.* IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2015. **23**(11): p. 2595-2605.
10. Sobe, U., A. Graupner, E. Böhme, A. Ripp, and M. Pronath. *Analog ip porting by topology conversion and optimization*. in *IP-ESC09 Conference*. 2009.
11. Fu, D.-S., Y.-Z. Chaung, Y.-H. Lin, and Y.-L. Li. *Topology-driven cell layout migration with collinear constraints*. in *Computer Design, 2009. ICCD 2009. IEEE International Conference on*. 2009. IEEE.
12. Kar, P.K. and S.K. Roy. *TECHMIG: A layout tool for technology migration*. in *VLSI Design, 1999. Proceedings. Twelfth International Conference On*. 1999. IEEE.
13. Diodato, P., A. Goksel, F. La Rocca, and W. Troutman. *Porting of a CMOS VLSI Chip from 2.5 Micron to 1.75 Micron Design Rules*. in *Solid-State Circuits Conference, 1984. ESSCIRC'84. Tenth European*. 1984. IEEE.
14. Johnson, N. *Agile hardware development–nonsense or necessity*. EE Times, 2013. https://www.eetimes.com/agile-hardware-development-nonsense-or-necessity/#.
15. Sangiovanni-Vincentelli, A., *The tides of EDA.* IEEE Design & Test of Computers, 2003. **20**(6): p. 59-75.
16. Bryant, R.E., K.-T. Cheng, A.B. Kahng, K. Keutzer, W. Maly, R. Newton, L. Pileggi, J.M. Rabaey, and A. Sangiovanni-Vincentelli, *Limitations and challenges of computer-aided design technology for CMOS VLSI.* Proceedings of the IEEE, 2001. **89**(3): p. 341-365.
17. López, R.C., F.V. Fernández, Ó. Guerra-Vinuesa, and Á. Rodríguez-Vázquez, *Reuse-based methodologies and tools in the design of analog and mixed-signal integrated circuits*2007: Springer Science & Business Media.
18. Macintosh, M., *Analog Hard IP made portable.* Chip Design Magazine, 2014.

19. Payne, D. *Three Steps for Custom IC Design Migration and Optimization*. 2016. https://www.semiwiki.com/forum/content/6140-three-steps-custom-ic-design-migration-optimization.html.
20. Dalbudak, E., El Guemhioui Karim, Lakhssassi Ahmed, *On the Use of Agile Methods for IC Design*, in *28th International Conference on Computer Applications in Industry and Engineering (CAINE 2015)*2015: San Diego, California.
21. Highsmith, J., A. *What is Agile Software Development?* CROSSTALK The Journal of Defense Software Engineering, 2002. http://www.adaptivesd.com/articles/cross_oct02.pdf.
22. Fowler, M. and J. Highsmith, *The agile manifesto*. Software Development, 2001. **9**(8): p. 28-35.

Breast Cancer Diagnosis in Histopathological Images Using ResNet-50 Convolutional Neural Network

Qasem Abu Al-Haija, Adeola Adebanjo
Department of Electrical and Computer Engineering, Tennessee State University,
3500 John A. Merritt Blvd, Nashville, TN 37209
Qabualha@Tnstate.edu, *Aadebanj@Tnstate.edu*,

***Abstract*— Breast cancer disease is the second most common world cause of cancer death in women. However, the early diagnostics and detection can provide a significant chance for correct treatment and survival. In this work, we propose an accurate and inclusive computational breast cancer diagnosis framework using ResNet-50 convolutional neural network to classify histopathological microscopy images. The proposed model employs transfer learning technique of the powerful ResNet-50 CNN pretrained on ImageNet to train and classify BreakHis dataset into benign or malignant. The simulation results showed that our proposed model achieves exceptional classification accuracy of 99% outperforming other compared models trained on the same dataset.**

***Keywords* — Breast Cancer, Histopathological Images, Medical image processing, ResNet-50, Convolutional Neural Network (CNN), Deep learning, Transfer Learning.**

I. INTRODUCTION

Cancer disease causes cells to divide uncontrollably in which abnormal cells grow and invade healthy cells in the body [1]. This can result in tumors, damage to the immune system, and other impairment that can be fatal. Indeed, the majority of cancer cases in the U.S. is breast cancer as reported by to the [2]. Breast Cancer Disease (BCD) is basically described as excessive or uncontrolled growth of breast tissues occurs. According to World Health Organization (WHO) [3], BCD is the second most common cause of death from cancer in women. However, BCD can be treated if it is detected at the earlier stages as reported by National Breast Cancer Foundation (NBCF) [4]. Therefore, BCD detection system is significantly on-demand to be addressed.

To cope with such medical image detection/classification task, deep learning (DL) has evolved as a subset of artificial intelligence that does its inferencing using deep neural networks by employing the artificial neural networks with several layers among the input layer and output layer. Indeed, image classification task using deep learning techniques has generated a lot of interest in recent years. ImageNet Large Scale Visual Recognition Challenge (ILSVRC) has been a genuine platform for researchers to showcase new ideas for classification. For example, in 2010, the use of GPUs, and a highly optimized implementation of 2D convolution was demonstrated as powerful enough to facilitate the training of large Convolutional Neural Networks (CNNs) [5].

CNNs have at least one Convolution layer, wherein instead of matrix multiplication, a convolution operation is performed on the input matrix in order to learn distinct low-level and high-level features of the image [6]. Deep CNNs are able to learn more features by increasing the depth of the network. However, increasing the depth of the network results in problems of vanishing gradients and degradation [7]. Thereafter, the continuous development in deep neural networks has enrich the AI field with the residual learning framework was presented in 2015 to ease the training of deep CNN networks [8]. This framework resulted in easier optimization of the network, and a higher accuracy. The network, later known as ResNet, was the basis of submissions to ILSVRC competition, where it won the first place on the task of ImageNet detection and ImageNet localization [9].

Residual neural networks (ResNet) address these challenges by introducing a "Residual block", which features a "skip connection", that adds the output from the previous layer to the layer ahead as illustrated in Fig.1. If x and $F(x)$ below do not have the same dimension, x is multiplied by a linear projection W to equalise the dimensions of the short-cut connection and the output layer. Empirical results from [9] demonstrate that the network is able to maintain stability even with far more layers than typical Convolutional Neural Networks.

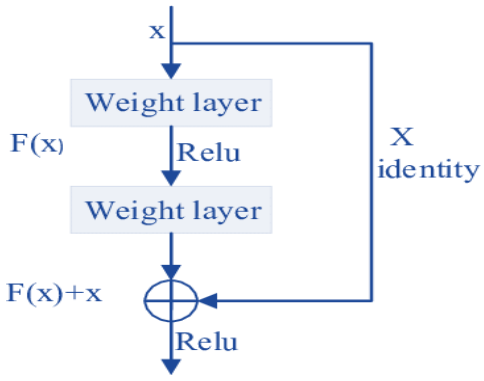


Fig. 1. Residual Network Building Block

In this paper, we propose to use a ResNet-50 (residual CNN with 50 layers deep) to produce classifications of histopathologic images to help providing an early detecting of BCD. The BCD dataset stores different kinds of breast cancer tissues and classifies them as either benign or malignant. While

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

the benign tissues are non-cancerous tumors, malignant tumors are cancerous, and can invade nearby tissue or spread to other parts of the body. Based on the collected images, a ResNet-50 network is trained using transfer learning. We show that the testing accuracy of our model is superior. In particular, the core contributions of the proposed work can be listed as follows:

- We provide a comprehensive efficient classification model that can classify the breast cancer imaging of BreakHis dataset into benign and malignant. Besides, the we present detailed preprocessing operations for the collected medical images prior to the use with deep learning algorithms.
- We employ the transfer learning technique for ResNet50 CNN that is pre-trained with ImageNet dataset to learn the new features for BreakHis dataset leveraging the power of free access GPU runtime provided by Google Co-laboratory.
- Extensive experimental findings are given to provide more insight into the proposed architecture and methodology. This includes simulation results related to the classification error and accuracy for training and testing as well as benchmarking of our results with existing related work.

The remainder of this paper is structured as follows: the next section, section II describes and discusses the system design modeling and architecture. Section III provides details about experimental environment, evaluation, and discussion. Finally, Section IV concludes the paper.

II. System Modeling Architecture

In this section, we describe our proposed system model which comprises four modules including: data collection module, data preprocessing module, feature learning module and data classification module. The complete system architecture showing all components is illustrated in Fig. 2. All subsystems (modules) are explained in the upcoming subsections.

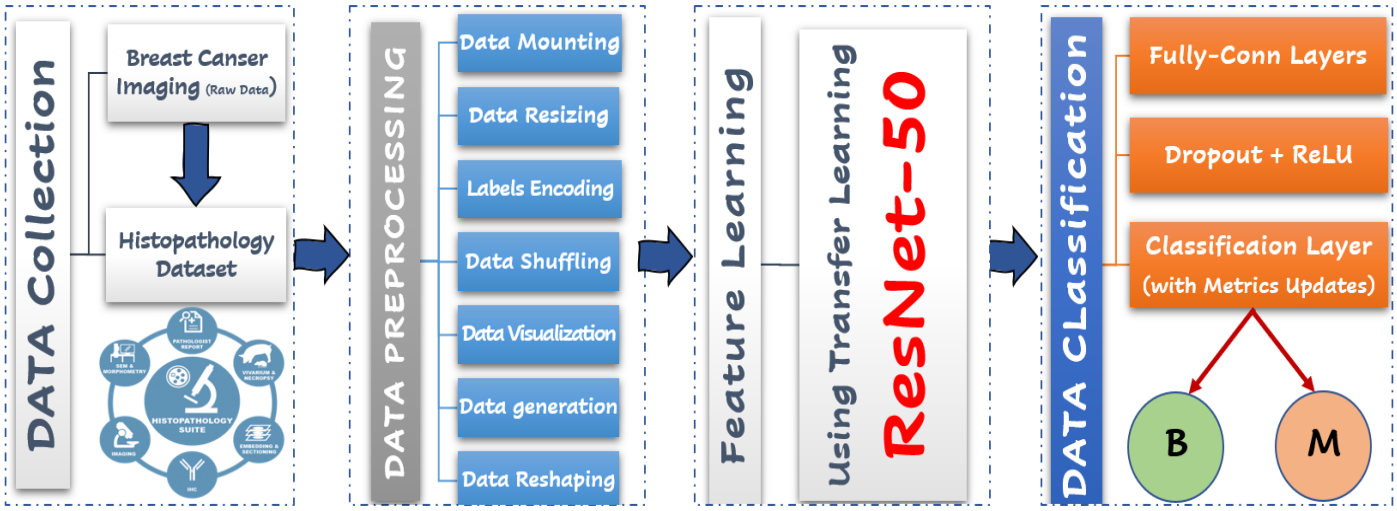


Fig. 2. Proposed System Architecture (Comprehensive data flow and system modules)

2.1 Data Collection Module

BCD data are available in the form of histopathological stained images [10] datasets such as WDBC breast cancer dataset [11], MITOS & ATYPIA-14 breast cancer dataset [12] and BreakHis-16 breast cancer dataset [13]. In this research, we are employing the BreakHis histopathology images breast cancer dataset to implement a deep detection system that can be used to provide binary classification of the BC image-data into *benign* (B) or *malignant* (M). Illustration samples for the benign and malignant images are provided in Fig. 3. BreakHis composed of RGB images (700 x 460 resolution) are taken by an accurate system composed of high-resolution camera (Samsung SCC-131AN) coupled with a 3.3x microscopic unit (Olympus BX-50). Images are captured in four different magnification levels that are equivalently distributed (~25% for each level) as provided in Table 2.

TABLE I. Imgaes' Distribution In The BreakHis Dataset

Category	Magnification level				Total
	40X	100X	200X	400X	
Benign	652	644	623	588	2480
Malignant	1370	1437	1390	1232	5429
Total	1995	2081	2013	1820	7909

Besides, there are many reasons to for this selection including:

- A comprehensive dataset with 7909 biopsy histopathology images breast cancer including both benign and malignant images acquired on 82 patients, publicly available.
- A Recent dataset associated with automated classification tasks, published by IEEE Transactions on Biomedical Engineering, 2016.
- A common histopathology dataset, in which its automated classification system is would be very valuable computer-aided diagnosis tool for clinician, if developed with high accuracy classification by employing CNN techniques.

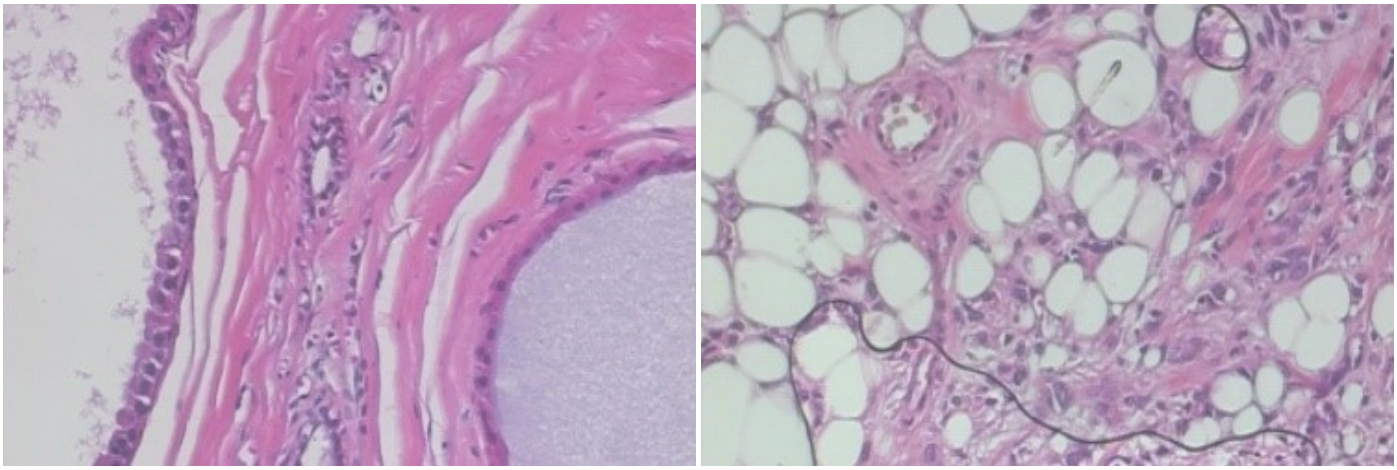


Fig. 3. Sample Images (a) Benign (b) Malignant.

2.2 Data Preprocessing Module

Generally, data preprocessing term belongs to all the conversions performed on the raw data prior to be processed by the deep learning module. For instance, training a convolutional neural network on raw images will probably lead to bad classification performances [14]. In this work, the collected dataset has passed through seven preprocessing operations (as shown in the second stage of Fig. 2) before it is fed to the next ResNet module, these operations are:

Data Mounting: This stage is used to mount Google Drive Account (DGA) as a virtual drive, just like a USB Drive on windows OS so you can browse and access your Drive from Google Co-laboratory. Therefore, we have uploaded our BreakHis dataset into a folder (named as BCDataset) on Google drive and then accessed the data into Co-Lab using $Python/glop$ library that enable the reading of dataset from external folders and $Python/pandas$ library that provides data manipulation including data framing, data reading and writing between in-memory data structures and different formats [15].

Data Resizing: This stage is necessary to remove redundancy from the input data which only contributes to the computational complexity of the network without providing any significant improvements in the result. This is achieved with the help of $Python/keras.preprocessing$ library. As a result of traying several image dimensions, we end up with image dimension of $144\ x\ 96$ which minimizes the size of image dimensions by a factor of 5 while maintaining image readability with efficient computational complexity.

Data Encoding: This stage is used to convert categorical data (textual data) into numerical values in which our deep learning predictive models can understand. In this step, we have employed the label encoding technique to converting each value in a the category column into a number, that is it, the category 'Benign' has given the value '0' while the category 'Malignant' has given the value '1'. This is achieved with the help of $Python/sklearn.preprocessing$ library.

Data Shuffling: This stage is used to redistribute the data samples of the training dataset to ensure that each data sample creates an "independent" change on the model, without being biased by the same points before them. This is achieved with the help of $Python/keras.preprocessing$ library.

Data Visualization: This stage is necessary to sample and examine the input data to ensure the readability of the input images by plotting few random samples of the training dataset via 2D representation with the new image dimensions. This is achieved with the help of both $Python/tensorflow$ and $Python/numpy$ libraries.

Data Generation: This stage is used to generate batches of tensor image data with real-time data augmentation. The data will be looped over (in batches) for both training and testing. Also, batch normalization is performed at this stage with image plotting for sample normalized images along with encoded labels. This is achieved with the help of both $Python/tensorflow$, $Python/keras.preprocessing$ and $Python/matplotlib$ libraries.

Data Reshaping: This stage is used to customize the input layer of ResNet-50 to accommodate the input shape for our preprocessed dataset ($Img_{Width} = 144$, $Img_{Height} = 96, No_{Channels} = 3$). This is achieved with the help of $Python/keras.preprocessing$ library.

2.3 Feature Learning Module

In machine learning, feature learning (FL) is a set of techniques that allows a system to automatically discover the representations needed for feature detection, prediction, or classification from the preprocessed dataset [16]. This allows a machine to learn the features and use them to perform a specific task such as classification or prediction. In deep learning, the feature learning can be accomplished by developing a complete convolutional neural network (CNN) to

train and test the set of images or by customizing a pretrained CNN in the classification/ prediction for the new images-set. The later technique is called Transfer learning. The idea of both techniques is illustrated in Fig.4. According to the figure, with transfer learning, you use the convolutional base (green module in the figure) and only re-train the classifier to your dataset (pink module).

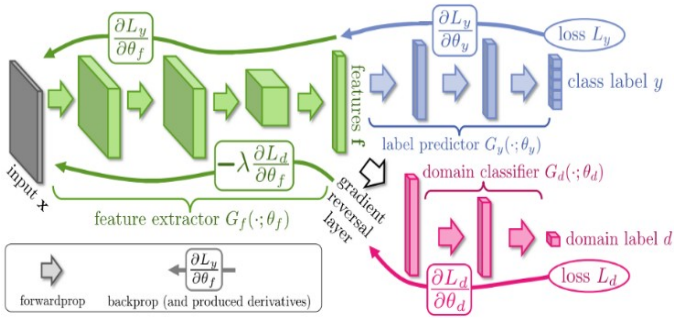


Fig. 4. Illustration of no-transfer learning vs transfer learning CNN [17].

Transfer learning (TL) is usually employed in the applications of DL which enable you to utilize a pretrained network to perform new prediction/classification tasks. This, indeed, require fine-tuning the learning parameters of the utilized pretrained network with randomly initialized weights to accommodate the new learning tasks. TL usually provides much faster and easier learning/training than training the network from scratch. As reported in [18], transfer learning is an optimization, a shortcut to saving time or getting better performance. This is illustrated in Fig. 5 that analyze the training performance of CNN employing transfer learning vs CNN with no transfer learning.

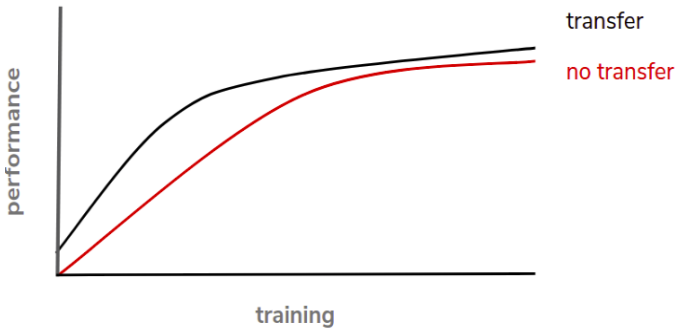


Fig. 5. Training performance of no-transfer learning CNN vs transfer learning CNN **[18]**.

In this module, we are utilizing the transfer learning technique to retrain the powerful ResNet-50 CNN with Fine-tuning for the network parameters and hyperparameters. This is accomplished by creating a model of ResNet-50 with pretrained parameters (weights) from *imagenet dataset* [19] after the preprocessing of the collected dataset (histopathologic images). This is achieved with the help of both *Python/ keras.applications.resnet50* and *Python/keras.models*. Illustration of this module can be depicted from Fig.6.

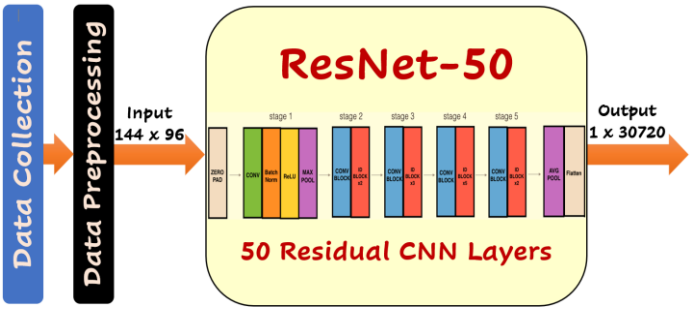


Fig. 6. Feature Learning Module Using ResNet-50 Transfer Learning

2.4 Data Classification Module

Data classification is an essential feature to separate large datasets into classes for the purpose of decision making, pattern recognition and others [20]. A classification layer makes use of fully connected layer and computes the cross-entropy loss for multi-class classification problems with mutually exclusive classes. This is achieved using *Paython/ keras.layers*, *Paython/keras.models* and *Paython/ keras/optimizers*. Illustration of this module can be depicted from Fig.7.

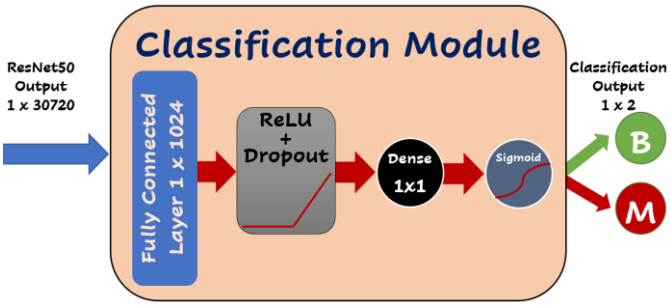


Fig. 7. Data Classification Module

According to the figure, the module receives the features from ResNet-50 and pass through a fully connected (FC) layer composed of 1024 neurons configured a 40% of dropout to prevent over-fitting [21]. After that, the units have been activated with rectification function namely known as *ReLU*. *ReLU* function is $MAX\,(X, 0)$ that sets all negative values in the matrix X to zero while all other values are kept constant. The reason of using *ReLU* is that training a deep network with *ReLU* tended to converge much more quickly and reliably than training a deep network with sigmoid activation [22]. Finally, to provide the probabilities for the classes, the output layer was composed of one neuron unit configured with Sigmoid function (Binary classifier). *Sigmoid* is mathematical function that takes as input a vector of K real numbers and normalizes it into a probability distribution consisting of two probabilities (e.g. Benign vs Malignant) [23]. *Sigmoid* function is defined as follows:

$$S(x)_i = \frac{1}{1+e^{-x}} = \frac{e^x}{e^x+1}, \quad S: \mathbb{R}^k \mapsto \mathbb{R}^k$$

$$for\ i\ =\ 1, 2\ \ldots, K\ and\ x\ =\ (x_1, x_1, \ldots, x_K) \in \mathbb{R}^k$$

Besides, to calculate the loss for training and testing we have employed *Mean Squared Error (MSE) loss*. MSE Loss [24] is calculated as the average of the squared differences between the predicted and actual values. The result is always positive regardless of the sign of the predicted and actual values and a perfect value is 0.05 MSE function is defined as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

Where: n is the number of samples, Y_i is the original data sample, and $\hat{Y}_i$ is the predicted data sample

Finally, to finalize the compilation of *Keras model*, we have utilized Root Mean Square Propagation optimizer (RMSprop) [25]. *RMSprop* utilizes the magnitude of recent gradients to normalize the gradients. In *RMSprop*, we divide the learning rate for a weight by a running average of the magnitudes of recent gradients for that weight. That is, it, keep a moving average of the squared gradient for each weight. The update is done separately for each parameter as follows:

$$For\ each\ Parameter\ w^j$$
$$(j\ subscript\ dropped\ for\ clarity)$$
$$\nu_t = \rho\nu_{t-1} + (1-\rho) * g_t^2$$
$$\Delta\omega_t = -\frac{\eta}{\sqrt{\nu_t + \epsilon}} * g_t$$
$$\omega_{t+1} = \omega_t + \Delta\omega_t$$

η : *Initial Learning rate*
ν_t : *Exponential Average of squares of gradients*
g_t : *Gradient at time t along* ω^j

III. EXPERIMENTAL ENVIRONMENT AND EVALUATION

To accomplish this proposed classification task, we have utilized different development tools and packages including *Python* 3.7 language along with *TensorFlow* 2.0 package and other aforementioned libraries, *Google Colaboratory* development environment leveraging the power of free access *GPU runtime* and *Google Drive* for storing and accessing the dataset. Besides, the experimental setup for training/testing model has been configured as follows: 75% of the dataset used for training (i.e., ~6000 images, here we used all images with 100X, 200X, 400X magnifications from both classes for training), 25% of the dataset used for testing (i.e., ~2000 images, here we used all images with 40X magnification from both classes for testing), the number of epochs = 200, the number of steps per epoch (iterations)=200, the number of verbose=1 (i.e., one progress bar line per epoch), and the *batch_size* = 32. This configurations have been achieved using *Python/model.fit* & *Python/matplotlib* libraries.

Also, the plot for loss function comparing the behavior of training loss and testing loss obtained during the training process is presented in Fig. 8. It can be clearly seen, both losses are systematically decreasing while training proceeds with faster threshold obtained for the training loss after only 25 epochs. However, the testing loss has saturated after almost 125 epochs of training process with less than 0.05 of MSE. This difference in saturation levels and threshold of training loss and testing loss is permitted to avoid underfitting or overfitting.

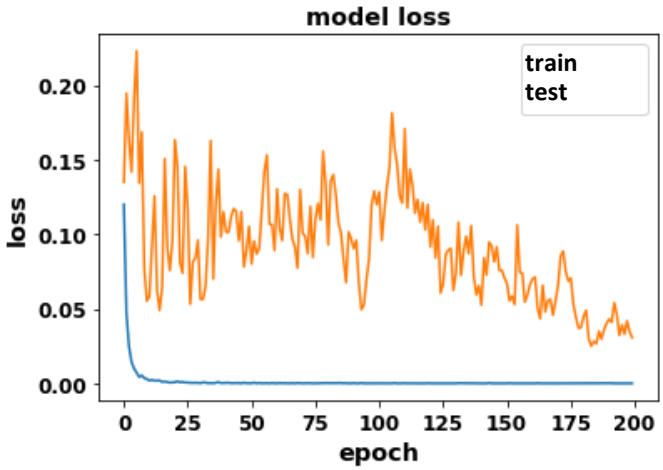


Fig. 8. Training/Testing Losses vs. number of epochs

Moreover, the plot for accuracy metric comparing the performance of training accuracy and testing accuracy obtained during the training process is given in Fig. 9. According to the figure, both accuracy curves are steadily increasing while training proceeds with faster ceiling level obtained for training accuracy after which recorded almost 100% only after 25 epochs. While the testing accuracy level was fluctuating between 98% and 99.8% after 100 epochs recoding an average testing accuracy of 99% of overall testing accuracy.

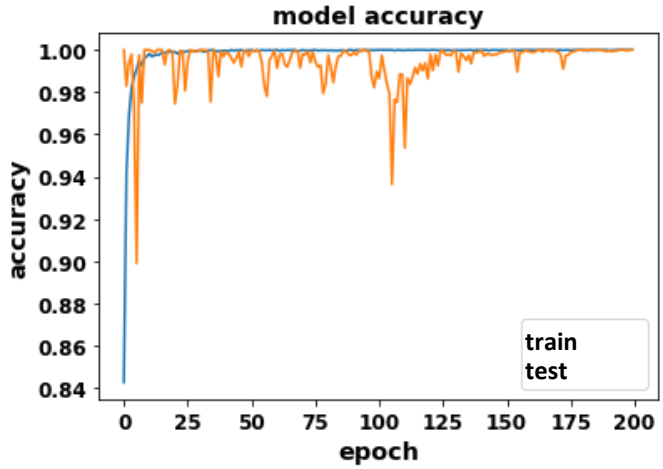


Fig. 9. Training/Testing Accuracies vs. number of epochs

Finally, even though the exiting state-of-art researches for classifying the histopathological images dataset use different network configurations, learning policies, programming techniques, and computing platforms, however, we still can compare the classification system performance in terms of training and testing accuracy

metrics. Therefore, for better readability, we summarize the time accuracy metrics for related state-of-art research's in the following table, Table I, in chronological order. According to the comparison of the table, it can be seen that our proposed model has recorded an attractive result in terms of both training and testing accuracy showing superiority over all other compared methods.

TABLE. II. COMPARISON WITH STATE OF ART METHODS: ACCURACY METRIC

Research	Method	Accuracy
Gour et. al., 2020 [26]	Customized ResHist [152- Residual Learning-CNN]	84.34%
Gupta et. al., 2020 [27]	Hybrid CNN [Employed Several Pre-Trained CNNs)	93.27%
Dabeer et. al., 2019 [28]	Customized LeNet-5 CNN	93.45%
Sagar et. al., 2019 [29]	Pre-Trained DenseNet201CNN	98.30%
Kassani et. al., 2019 [30]	Pre-Trained Combined CNNs [DenseNet201+ VGG19 + MobileNetV2]	98.13%
Gandomkar et al., 2018 [31]	Pre-Trained ResNet-152 CNN	98.77%
Adeshina et. al., 2018, [32]	New Deep CNN/14 Layers (DCNN-14)	91.5%
Han et al., 2017 [33]	New Class Structure-Based Deep CNN (CSDCNN)	93.20%
Sun, et. al, 2017 [34]	Pre-Trained GoogLeNet CNN	95.00%
Spanhol et al. 2016 [35]	Pre-Trained AlexNet CNN	84.60%
Proposed Method	Pre-Trained ResNet-50 CNN	**99.10 %**

IV. CONCLUSIONS

An efficient model for classifying the stained histological breast cancer images with high level of classification accuracy. To increase the robustness of the classifier, we employed the transfer learning of the powerful ResNet-50 CNN pretrained on ImageNet. The developed model makes use of *BreakHis* dataset with 75% of the images used for training and 25% used for testing. Indeed, the proposed work provides a comprehensive model for medical image processing/classification from input layer to the output layer. Eventually, to our knowledge, the reported results are superior to the automated analysis of breast cancer images reported in literature [26-35].

REFERENCES

[1] R. Nall, Y. Ranchod, What to know about cancer, Medically reviewed by Yamini Ranchod, Medical News Today (MNT), Jan, 6, 2020.

[2] NCI, Breast Cancer—Patient Version, National Cancer Institute (NCI), 2020.

[3] WHO. (2018). Data of breast cancer sufferers in the world. World Health Organization (WHO), Jan. 3. 2019.

[4] NDCF, Learn about breast cancer, 2019 National Breast Cancer Foundation.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet classification with deep convolutional neural networks. Commun. ACM 60 (6), pp. 84–90, 2017.

[6] S. Saha. A Comprehensive Guide to Convolutional Neural Networks-ELI5 way. Towards Data Science, A Medium publication sharing concepts, ideas, and codes, 2018.

[7] S. Jia. Vanishing Gradient vs Degradation. Towards Data Science, A Medium publication sharing concepts, ideas, and codes, Sep, 2018.

[8] C. Shorten. Introduction to ResNets. Towards Data Science, A Medium publication sharing concepts, ideas, and codes, Jan, 2019.

[9] K. He, X. Zhang, S. Ren, J. Sun. Deep Residual Learning for Image Recognition. Microsoft Research. 2015.

[10] J. D. Toro, et. al., Analysis of Histopathology Images. Chapter 10, Biomedical Texture Analysis Fundamentals, Tools and Challenges, Elsevier, 2017, Pages 281-314, 10.1016/B978-0-12-812133-7.00010-7.

[11] W. H. Wolberg, W. N. Street, O. L. Mangasarian, Breast cancer Wisconsin (diagnostic) data set, UCI Machine Learning Repository

[12] Lowe A KMe a, Grunkin M. Mitos atypia grand challenge 2014. 2014. https: //mitos-_atypia-_14.grand-_challenge.org/Dataset/

[13] Spanhol, F., Oliveira, L. S., Petitjean, C., Heutte, L. A Dataset for Breast Cancer Histopathological Image Classification, IEEE Transactions on Biomedical Engineering (TBME), 63(7):1455-1462, 2016. DOI: 10.1109/TBME.2015.2496264

[14] K. K. Pal and K. S. Sudeep. Preprocessing for image classification by convolutional neural networks. IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, 2016, pp. 1778-1781.

[15] W. McKinney. pandas: powerful Python data analysis toolkit Release 1.0.3. By Pandas Development Team.

[16] Y. Bengio; A. Courville; P. Vincent. Representation Learning: A Review and New Perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence. 35 (8): 1798–1828, 2013. arXiv:1206.5538. doi:10.1109/tpami.2013.50. PMID 23787338.

[17] Ganin, Y., & Lempitsky, V. Unsupervised Domain Adaptation by Backpropagation. ACM 32nd International Conference on Machine Learning, vol. 37, 2015.

[18] R.F. Lopes. Wild Data Part 3: Transfer Learning. Stratio Big Data Inc, 2018.

[19] K. He, X. Zhang, S. Ren and J. Sun . Deep Residual Learning for Image Recognition. arXiv:1512.03385v1, 2015.

[20] O. Knocklein. Classification Using Neural Networks. Medium: Toward Data Science. Jun. 2019.

[21] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. J. Mach. Learn. Res., 15(1):1929–1958, Jan. 2014.

[22] J. Brownlee. A Gentle Introduction to the Rectified Linear Unit (ReLU). Machine Learning Mastery Inc., Deep Learning Performance, Jan, 2019.

[23] Fei-Fei Li, "CS231n: Convolutional Neural Networks for Visual Recognition", Computer Science, Stanford University, 2019.

[24] M. Binieli .Machine learning: an introduction to mean squared error and regression lines. #MATHEMATICS, Free Code Camp, 16 OCT 2018

[25] A. C. Wilson, R. Roelofs, M. Stern, N. Srebro, B. Recht . The Marginal Value of Adaptive Gradient Methods in Machine Learning. Advances in Neural Information Processing Systems 30 (NIPS 2017), Pp. 4148—4158, 2017.

[26] M. Gour, S. Jain, T. S. Kumar, Residual learning based CNN for breast cancer histopathological image classification. Int J Imaging Syst Technol. 2020; 30: 621– 635. https://doi.org/10.1002/ima.22403

[27] K. Gupta, N. Chawla, Analysis of Histopathological Images for Prediction of Breast Cancer Using Traditional Classifiers with Pre-Trained CNN, Procedia Computer Science,Volume 167, 2020, Pages 878-889, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.427.

[28] S. Dabeer, M. M. Khan, S. Islam. Cancer diagnosis in histopathological image: CNN based approach. Informatics in Medicine Unlocked, Elsevier, Vol.16, 2019, 100231.

[29] A. Sagar. Convolutional Neural Network for Breast Cancer Classification: Deep Learning for solving the most commonly diagnosed cancer in women. Towards Data Science, A Medium publication sharing concepts, ideas, and codes, 2019.

[30] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R.,Deters. Classification of histopathological biopsy images using ensemble of deep learning networks. ACM 29th Annual International Conference on Computer Science and Software Engineering (CASCON '19). IBM Corp., USA, 92–99, 2019.

[31] Z. Gandomkar, P. C. Brennan, and C. Mello-Thoms. MuDeRN: Multi-category classification of breast histopathological image using deep residual networks. Artificial Intelligence in Medicine 88, pp.14 – 24, 2018.

[32] Adeshina SA, Adedigba AP, Adeniyi AA, Aibinu AM. Breast cancer histopathology image classification with deep convolutional neural networks. In: 2018 14th international conference on electronics computer and computation (ICECCO). IEEE; 2018. p. 206–12.

[33] Z. Han, B, Wei, Y. Zheng, Y. Yin, K. Li, and S. Li. Breast cancer multi-classification from histopathological images with structured deep learning model. Scientific reports 7, 1 (4172), 2017.

[34] J. Sun and A. Binder, "Comparison of deep learning architectures for H&E histopathology images," 2017 IEEE Conference on Big Data and Analytics (ICBDA), Kuching, 2017, pp. 43-48.

[35] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte. Breast cancer histopathological image classification using convolutional neural networks. International joint conference on neural networks (IJCNN). IEEE, 2560—2567, 2016.

IoT-based Recommendation Systems – An Overview

Dina Nawara, Rasha Kashef
Electrical, Computer and Biomedical Engineering Department
Ryerson Umiversity,
Toronto, Ontario, Canada
{dina.nawara, rkashef}@ryerson.ca

***Abstract*—Internet of Things (IoT) has emerged in many industries, such as health care, transportation, agriculture, manufacturing, smart homes, to name a few. It paves the path for massive applications on the user level to enhance the quality of life or service, and on the decision-makers' level to provide a sustainable increase in revenue. IoT principally connects different physical objects (e.g., sensors) and enables them to communicate, collect, and share data. In the Era of IoT, Recommendation systems provide personalized recommendations based on the user's historical datasets collected from the IoT devices. These recommendations enable an efficient decision-making process by suggesting relevant products, resources, and information. This paper provides an overview of various multi-layers IoT architectures, and IoT-based recommendation systems with an emphasis on their advantages, disadvantages, application domains, and validation metrics for quality assessment.**

Keywords—IoT, Architecture, Recommendation systems, Evaluation Metrics.

I. INTRODUCTION

The Internet of Things (IoT) has gained the attention of researchers as well as major industries and markets in the last few decades. IoT connects heterogeneous physical devices to collect data using (sensors, actuators, etc.). It uses the collected data in many applications to enhance the outcomes of a decision making process in various applications, such as Monitoring systems, healthcare, transportations, data storage, smart-homes. With the abundant size of IoT-related datasets, recommendation systems come to provide suggestions to users and enterprises or organizations. These recommendations are tailored to the user's preferences, thus enhancing the user's experience as well as the enterprises goal. Current recommendation systems can be classified as content-based filtering, collaborative-based filtering, knowledge-based, demographic-based, and hybrid recommendation systems [1], [2].

Content-based filtering system [1] uses content information of the items to build the a recommendation model. It recommends items based on every user's profile and the history of decisions and preferences. It uses different models to find similarity between items. It is commonly used in videos, movies, and music recommendations. It can adjust to new profiles quickly. However, it has some limitations, since it requires a very well organized user's profile to provide a recommendation [3]. In contrast, Collaborative recommenders work as the word-of-mouth theory. They recommend information to the user's level considering similar preferences of other users [4]. A group of similar users is called neighbors, such that a neighborhood is a set of like-minded users. Knowledge-based recommenders use an explicit knowledge about the users' preferences, items' classifications and user-item matching relations. This recommendation technique does not face the cold start problem that collaborative filtering recommenders face. However, it requires explicit recommendation knowledge definition [5]. Demographic-based recommender systems classify users according to their demographic features, such as age, occupation, and gender. Security and privacy issues are considered to be the main disadvantages of these models. Hybrid filtering systems combine two or more filtering techniques, and their main function is to mitigate the constraints faced by each filtering approach when used solely. For example, combining content-based filtering and collaborative-based filtering systems increase the recommendations' precision [6].

Recommendation systems face challenges and constraints, such as (data scalability, data sparsity, and cold start).

- *Cold start* [7] rises when there is no enough history for the new user to find a match or to recommend a relevant value. This problem can be avoided using hybrid systems.
- *Data sparsity* occurs when the user-rating matrix is sparse, and it is difficult to find similar users, hence the efficiency of the recommender is impacted.
- *Data Scalability is* the increase of time and computational cost needed to process massive data [8].

In this paper, we review some IoT-based recommendation systems by highlighting their purposes, advantages, disadvantages, challenges, data sources, and methods. In this paper, section 2 introduces the concept of IoT and its architecture. Section 3 provides a literature of the existing proposed recommendation systems and its IoT applications, and section 4 presents evaluation metrics. Conclusion and future work are discussed in section 5.

II. INTERNET OF THINGS – IOT

IoT provides opportunities in industries as transportation, wearables devices, smart homes, healthcare [9], [10] to name a few. (Fig.1) for tracking, monitoring, and operating remote devices or things [11].

978-1-7281-9615-2/20/$31.00 ©2020 Crown

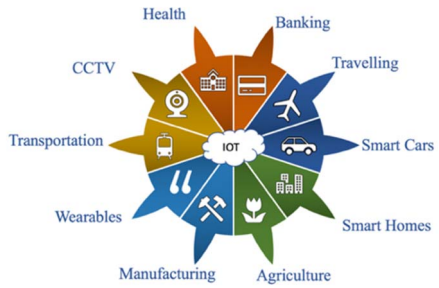


Fig. 1. IoT Market Applications

IoT is an infrastructure with standardized and self-configuring capabilities with interoperable communication protocols, where different devices "Things" have unique identifiers (UID), and they are integrated into the network [12]. The number of IoT connected devices is increasing dramatically; according to the connected devices statistics [13], there are 20 Billion connected units in 2019. IoT applications are classified as customer-based applications, such as (wearables, smart homes), infrastructure-based applications, like (smart environments and smart grid), and industrial-based applications, e.g. (smart retail, smart supply-chain) [14]. IoT has also been involved in many energy management related solutions aimed to benefit both consumers and service providers and increase energy consumption efficiency [15]. This can be achieved by optimizing the usage with activating services like demand and response based on the data collected from IoT home power sensors/monitors [16]. IoT based smart grids empowered the consumers and enabled them to better control their consumption, and manage the "things/devices" batteries' lifetime [17], [18], [19].

IoT is a multi-layered architecture that is proposed to meet business requirements and to ensure the interoperability of heterogeneous devices. These devices are connected in a heterogeneous mode, and they collect, process data and communicate this data in real-time [20]. There is no one single architecture model that is universally used. However, there are some proposed models with a different number of layers, for example (Three-layers architecture, Four-layers architecture, Five- layers architectures) [12], [21], [22], [23], Fig. 3 presents the different IoT architecture models, while Fig.4 demonstrates the basic Three-Layer model.

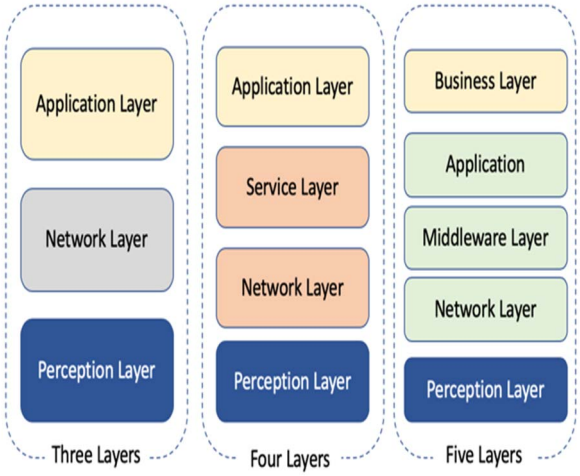


Fig. 3. IoT Architecture Models

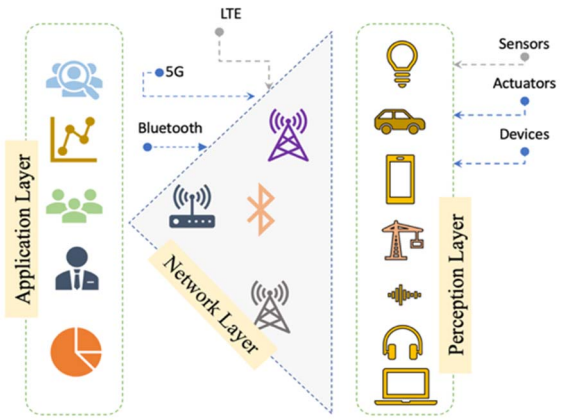


Fig.4. High-Level demonstration for the Three-Layers IoT Architecture

The Three-layer architecture is the basic model. It was implemented in many applications such as (smartphone, power management, and E-health applications) [24], and the other two models are then introduced to meet the operational and functional complexity of IoT. Some of these architecture layers are defined next.

- The *Perception Layer* has the interactions of physical sensors, actuators, identifiers, etc. Those physical devices have Radio Frequency Identification (RFID), which are used to detect and tag the objects/devices [25]. Perception Layer is responsible for collecting, measuring and controlling the data, then transferring them to the network layer.
- The *Network Layer* is also known as the transmission layer and is responsible for routing the data using different transmission technologies, like (Bluetooth, WiFi, LTE, 5G, ZigBee).
- The *Service Layer* provides the required services and makes decisions based on the received data.
- The *Middleware Layer* contains diverse and massive data coming from the network layer for the analysis. Programmers' role can be adopted in this layer to deploy different technologies like cloud computing, big data, etc.
- The *Application Layer* contains the services requested by the customers—applications like data storage, devices' behavior analysis, smart homes, smart healthcare systems, etc.
- The *Business Layer* manages the whole IoT model. It generates flowcharts, graphs, and statistics based on the data received from the Application Layer. It also measures every layer's output against the expected outputs.

III. IoT-based Recommenders

The volume of collected and produced data has eventually increased in the Era of IoT. Recommendation systems are deployed in different IoT domains, such as smart homes [26], healthcare [27] [28], agriculture [29], and intelligent transportation [30] [31] to create meaningful information or patterns that lead to reasonable decisions. In the following, we provide an overview of the application of recommendation systems in the IoT environment with focus on three main applications as healthcare, E-commerce, and smart transportation.

A. IoT in Healthcare – Recommendation systems

Smart health applications have attracted researchers in the last years since users started to use smart devices to share information with other devices and networks. Some researchers introduced recommendation systems to provide health services and improve the quality of living, such as remote diagnostics, remote provision of care, distant medical reports, etc. In [32] the importance of smart healthcare systems and how patients can proactively contribute to those systems is called "patient-centric solutions". They have used the data provided by patients' preferences, patients' health conditions, and the city routing information to create a recommendation system that recommends the city routes that work best for the citizens. In [33] a novel recommender system based on IoT applications is developed to tackle age-related health issues. Authors have collected user information using IoT technology and then recommended quality life solutions to the patients (step counter, wristband, virtual Nurse, etc.). They used supervised and semi-supervised algorithms and applied them to data collected from sensors measuring blood pressure, thermometers, heart rate, etc. In [34], they have developed an IoT-based recommendation system, which diagnoses cardiac disease symptoms and produces dietary plan recommendations. The authors used biosensors to collect the needed information and evaluated their recommender results using precision, recall, and MAE. In [35], recommenders like ProTrip are proposed with new functionality that helps travelers with their diet's chronic diseases by introducing health-centric e-tourism recommendations. The authors tailored the application interface to meet the needed data representation according to the customers' profiles. In [36], an IOT-based recommender system provides diet plans to patients using Type-2 Fuzzy logic ontology. The authors used wearable sensors to extract information about the patients' health conditions. Recommendation systems aim to benefit users and individuals in all aspects, including Well-being. Since there are many wearable devices that support day-to-day fitness, nutrition, and physiological health, these devices are producing signals that are being collected and analyzed to understand the physical and emotional states of the users. [37] designs a recommendation system that promotes user happiness, and they called it “IAMHAPPY”. They built a web-based repository for the emotional states, and they recommended naturopathy recommendations to ease the minor discomforts that users face. In [38] the researchers tackled the depression conditions and developed a mobile application to collect emotional-related data for users. They used 1047 volunteers’ data to train their model using supervised algorithms, and then they designed a recommendation system to recommend personalized solutions for the users. Table 1 provides a comparison between some healthcare recommender systems.

B. IoT in E-Commerce – Recommendation Systems

Nowadays, everyone uses smart devices in their day to day life (smartwatch, smartphone, smart tooth-brush, smart washer, etc.). Utilizing the collected e-business data, plenty of recommenders in E-commerce are provided. In [39], a mobile application to identify the list of smart physical objects that customers use is proposed, then the inventories are used to recommend further users’ needs. In [40], a recommendation system that recommends IoT devices to the users based on their interests is introduced. They used the similarity between users, services, and devices to recommend IoT objects accordingly using data sets from IoT providers (Telus, BlueRover, and Libelium). They defined relationships between the three components of their recommender (devices, users, and services), and evaluated their results using Precision, Recall, and root mean squared error (RMSE). The authors in [41] focused their studies on developing an IoT platform for real-time recommenders used for the tourism domain. They designed an IoT platform that recommends new items using geofencing as a user context, aimed to develop unique experience and exposure that connects the tourists with their surroundings. A novel Tourism recommender was also developed in [42] using route mining algorithms and travel behavioral data. The authors collected travel behavior information by operating tourists’ smartphones and based on the tourists’ behavior patterns. They recommended travel routes based on the travelers’ profiles. Authors chose a candidate like a park or a museum with sightseeing areas, and then, they installed Bluetooth low energy devices to locate the tourists. In [43], authors tackled problems in recommender systems such as the data sparsity, and they developed recommenders by adopting Word2Vec, which can be used to extract information and semantics from users’ comments ratings. The recommendation system relies detecting the feedback from users’ comments using semantic analysis and reducing the dimensionality to increase accuracy and speed. Some of the state-of-art E-commerce recommendation systems are shown in Table 2.

C. IoT in Smart Transportation – Recommendation Systems

In [44], a recommender system based on IOV (Internet of Vehicles) is proposed to adopt fog computing and use content-based filtering to recommend places. They adopted a fog computing environment aiming to provide users with the needed recommendations in terms of low latency and high accuracy. In [45], a Mobility Recommender System is developed to help users with their mobility by providing an effective End-to-End solution that matches their needs and aligns with the city regulations. In [46], the authors reused CityPulse framework that supports smart city services, to develop a proof of concept for a public transportation improvement planner. [47] a parking violation management recommender was designed, aiming to increase the officers’ productivity by recommending the least cost routes to find the parking violations. Also, in [48] an android mobile application for suggesting available parking slots was developed, taking different contextual attributes such as (parking facilities, walking distance, weather conditions, etc.). The authors utilized Context-as-a-Service (CoaaS) framework and evaluated their model using precision/recall. In [49], a parking conflicts’ solution was proposed, by providing a periodical recommendation to the nearest available parking slot, so that the users can drive their cars to it, quickly. The authors considered users’ preferences such as (preferred zone, and shopping zone). Table 3 shows a comparison between various smart transportation recommenders.

D. Dicussion

We can observe the following discussion points from the summarization provided in Tables 1-3:

- There is a research need to address user privacy and protecting users’ data from any malicious activity in the current recommendation systems.
- Most of the used datasets are simulation data.
- There are a few research studies conducted in the area of well-being and mental health-related personalized recommendation.

Table 1: IoT-based HealthCare Recommendation systems

Paper	Recommendation approach	Recommendation algorithm	Similarity measures	Dataset	Recommendation results	Evaluation
[32]	Collaborative filtering- Memory based	K- nearest neighbor	Euclidean distance	Route database from city of Tarragona in Barcelona, the other data were simulated	Smart City routes for citizens/patients based on their preferences	MAE
[33]	Collaborative Filtering	K- nearest neighbor	Simplified proposed similarity	Dataset from well-established m-health application was used	Life solutions/devices like (step counters, wristbands)	–
	Content Based filtering	–	Jaccard Similarity		Virtual Nurse	
[34]	Hybrid filtering system	–	–	Well-renowned hospital	1. classification to cardiovascular disease based on the data collected from bio sensors 2. dietary plans recommendations based on gender and age.	Precision, Recall and MAE
[35]	Hybrid filtering system	–	–	1. Climate-Based dataset 2. Food information dataset 3. User Dataset	dietary plans to the travelers who have chronic diseases	Precision, and Recall
[36]	Fuzzy ontology	Fuzzy Ontology	–	Dataset for drugs information	Recommending drugs and diets to patients with certain chronic diseases based on the health condition	Precision, and Recall
[37]	–	–	–	Simulation	recommend naturopathy remedies	Users' feedback
[38], [39]	Collaborative-based	–	–	Volunteer data (Simulation)	Personalized recommendation for depression's solutions	Users' feedback

Table 2 : IoT-based E-Commerce Recommendation Systems

Paper	Recommendation approach	Recommendation algorithm	Similarity measures	Dataset	Recommendation results	Evaluation
[40], [41]	Collaborative Filtering	–	Pearson Correlation Coefficient	Gathering data from companies like (Telus, Libelium, BlueRover)	Recommends IoT smart objects based on the users interests.	Precision, Recall and RMSE
[43]	Collaborative Filtering	–	–	Amazon dataset	recommends movies and TV shows based on the users' comments and ratings	MAE and RMSE
[42]	Content_Based Filtering	–	–	Simulation	Travel routes based on the travelers' profiles and behavior patterns	Rating questionnaires

Table 3 : IoT- based Smart Transportation Recommendation systems

Paper	Recommendation approach	Recommendation algorithm	Similarity measures	Dataset used	Recommendation results	Evaluation
[44]	Content_Based Filtering	Nearest neighbor	Euclidean distance	Simulation	Recommendations based on user score matrix	MAE and RMSE
[45]	–	–	–	Public transportation map, and city regulation	recommends suitable routes and parking spaces whether on or off - street	–
[49]	Content_Based Filtering	–	Minkowski distance	Simulation	Parking conflicts reduction by providing periodic recommendations for available parking slots	Precision, Recall and F measure
[48]	Context aware	–	Boolean, continues, semantic, similarity	Monash University parking data	Recommendation for Parking slots	Precision/Recall
[46]	Context aware	–	–	URBUS provider in Brasov, Romania	Planner for Public transportation	–
[47]	–	–	–	Simulation on parking sensors collected data in the City of Melbourne	Recommendations for the least cost paths to officers to find parking violations	Defined performances metrics such as (Average issued tickets ratio, number of issued tickets per hour)

IV. Recommender Systems Evaluation Metrics

As we rely on recommendation systems to provide essential decisions, evaluating the recommendation systems' performance is a critical task. The most systematic way for evaluation is to divide the dataset into train and test sets, build a model on the training set, and then use the test set for assessment. Metrics as MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), Precision, and Recall are commonly used to quantify the performance of IoT-based recommenders. Other evaluations may need the users' feedback in the form of questionnaires, or studies [50].

- Mean Absolute Error (MAE): It is an averaged measure between the recommender's predicted values/ratings, and the users' observed values/ratings [51]. The lower the MAE value is, the higher the recommender's correctness.

$$MAE = \frac{\sum_{i=1}^{n} |A_i - B_i|}{n} \quad (1)$$

A_i : is the predicted values/ratings

B_i: is the observed users' values/ratings

n: is the number of evaluated ratings

- Root Mean Square Error (RMSE): The RMSE measures the squared root of the deviations between residuals, where residuals are the difference between the predicted values/rating and the observed values/ratings [51] .

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (A_i - B_i)^2}{n}} \quad (2)$$

- Precision and Recall: Precision is the value of the used predicted recommendations by the users, to the total recommended values. Recall measures the used predicted recommendations to the total relevant values [6].

$$Precision = \frac{Used\ Recommended\ values}{Total\ recommended\ values} \quad (3)$$

$$Recall = \frac{Used\ Recommended\ values}{Total\ relavent\ values} \quad (4)$$

- F-measure is also called "F1-score", measures the harmonic mean of both precision and recall measures [7].

$$F\text{-}measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

V. Conclusion And Future Directions

Recommendation systems play an essential role in our daily lives. With the increase of IoT connected devices, the amount of data has dramatically increased. This has created enormous potential for recommendation systems in many industries and domains to provide tailored recommendations for users. These recommendations are used to improve patient's quality of health, customer shopping experience, or enhance the traveling plans, to name a few. In this paper, we have discussed the IoT context along with its architecture. We have also reviewed some developed recommendation systems in healthcare, e-commerce, and other domains of smart transportation in the IoT environment, along with its promises, pitfalls, and directions. Future work includes different domain applications, large-scale recommendation systems, and unsupervised machine learning recommenders.

References

[1] P. M. S. Dinesh Vijay Attarde, "Survey on Recommendation System using Data Mining and Clustering Techniques," *International Journal for Research in Engineering Application & Management (IJREAM),* vol. 3, no. 9, 2017.

[2] N. P. a. S. L. Varma, "Financial Planning Recommendation System Using Content-Based Collaborative and Demographic Filtering," in *dvances in Intelligent Systems and Computing*, Springer, Singapore, 2019.

[3] S. V. Nymphia Pereira, "Survey on Content Based Recommendation System," *International Journal of Computer Science and Information Technologies,,* vol. 7, no. 1, pp. 281-284, 2016.

[4] A. M. F. H. K. B. Basit Mehmood Khan, "Collaborative Filtering based Online Recommendation Systems: A Survey," in *International Conference on Information and Communication Technologies (ICICT)*, 2017.

[5] Z. N. John K. Tarus and G. Mustafa, "Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning," *Artificial Intelligence Review,* vol. 50, pp. 21-48, 2018.

[6] Y. F. B. O. F.O. Isinkaye a, "Recommendation systems: Principles, methods and evaluation," *Egyptian Informatics Journal,* vol. 16, pp. 261-273, 2015.

[7] H. G. P. B. Lipi Shah, "Survey on Recommendation System," *International Journal of Computer Applications,* vol. 137, no. 7, 2016.

[8] D. N. P. Utkarsh Gupta, "Recommender System Based on Hierarchical Clustering Algorithm Chameleon," in *IEEE International Advance Computing Conference (IACC)*, 2015.

[9] P. K. Deb, S. Misra, A. Mukherjee and A. Jamalipour, "SkopEdge: A Traffic-Aware Edge-Based Remote Auscultation Monitor," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, Dublin, Ireland, 2020.

[10] N. Pathak, A. Mukherjee and S. Misra, "Reconfigure and Reuse: Interoperable Wearables for Healthcare IoT," in *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, Toronto, ON, 2020.

[11] I. Y. Sajjad Hussain Shah, "A Survey: Internet of Things (IOT) Technologies, Applications and Challenges," in *the 4th IEEE International Conference on Smart Energy Grid Engineering*, 2016.

[12] W. H. S. L. Li Da Xu, "Internet of Things in Industries: A Survey," *IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS,,* vol. 10, no. 4, 2014.

[13] "Statista.com," Feb 2020. [Online]. Available: https://www.statista.com/statistics/764026/number-of-iot-devices-in-use-worldwide/.

[14] D. K. Sharu Bansal, "IoT Ecosystem: A Survey on Devices, Gateways, Operating Systems, Middleware and Communication," *International Journal of Wireless Information Networks,* vol. 27, pp. 340-364, 2020.

[15] I. A. Z. M. R. R. G. M. A. A.R. Al-Ali, "A Smart Home Energy Management System Using IoT and Big Data Analytics Approach," *IEEE Transactions on Consumer Electronics,,* vol. 63, no. 4, pp. 426 - 434, 2017.

[16] G. K. V. S. B.-C. W. Guneet Bedi, "Review of Internet of Things (IoT) in Electric Power and Energy Systems," *IEEE INTERNET OF THINGS JOURNAL,* vol. 5, no. 2, pp. 847 - 870, 2018.

[17] M. S. J. A. H. S. N. A. A. S. A. Y. Norizam Sulaiman, "Evaluation of Mobile Phone Wireless Charging System Using Solar and Inductive Coupling," in *International Conference on Advances in Information and Communication Technology*, 2016.

[18] "Fair Scheduling in Resonant Beam Charging for IoT Devices," *IEEE Internet of Things Journal ,* vol. 6, no. 1, pp. 641 - 653, 2018.

[19] A. K. H. E. Sihua Shao, "Pixelated VLC-backscattering for Self-charging Indoor IoT Devices," *IEEE Photonics Technology Letters,* vol. 29, no. 2, pp. 177-180, 2015.

[20] A. M. A. A.-A. X. D. I. A. M. G. Mohammed Ali Al-Garadi, "A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security," *IEEE Communications Surveys & Tutorials* , 2020.

[21] M. G. M. M. M. A. M. A. Ala Al-Fuqaha, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," *IEEE COMMUNICATION SURVEYS & TUTORIALS,,* vol. 17, no. 4, 2015.

[22] S. R. S. Pallavi Sethi, "Internet of Things: Architectures, Protocols, and Applications Review," *Journal of Electrical and Computer Engineering,* p. 25, 2017.

[23] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang and W. Zhao, "A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications," *IEEE Internet of Things Journal,* vol. 4, no. 5, pp. 125 - 1142, 2017.

[24] D. a. I. R. Mohammed Riyadh Abdmeziem, "Architecting the Internet of Things: State of the Art," in *Robots and Sensor Clouds. Studies in Systems, Decision and Control*, 2016, pp. 55-75.

[25] A. F. ,. M. C. A. S. Mohammed El-hajj, "A Survey of Internet of Things (IoT) Authentication Schemes," *sensors,* 2019.

[26] H. A. Aldhubaib and R. Kashef, "Optimizing the Utilization Rate for Electric Power Generation Systems: A Discrete-Event Simulation Model," in IEEE Access, vol. 8, pp. 82078-82084, 2020, doi: 10.1109/ACCESS.2020.2991362.

[27] R. Kashef and M. S. Kamel, "Cooperative Partitional-Divisive Clustering and Its Application in Gene Expression Analysis," 2007 IEEE 7th International Symposium on BioInformatics and BioEngineering, Boston, MA, 2007, pp. 116-122, doi: 10.1109/BIBE.2007.4375553.

[28] Shirazi, Z.A., de Souza, C.P., Kashef, R. and Rodrigues, F.F., 2020. Deep Learning in the Healthcare Industry: Theory and Applications. In Computational Intelligence and Soft Computing Applications in Healthcare Management Science (pp. 220-245). IGI Global.

[29] R. Kashef, "Adopting Big Data Analysis in the Agricultural Sector: Financial and Societal Impacts," *Internet of Things and Analytics for Agriculture,* vol. 2, pp. 131-154, 2020.

[30] F. Zantalis, G. Koulouras, S. Karabetsos and D. Kandris, "A Review of Machine Learning and IoT in Smart Transportation," *Future Internet,* vol. 11, no. 4, 2019.

[31] Z. Karami, R. Kashef, "Smart transportation planning: Data, models, and algorithms," *Transportation Engineering,* vol. 2, 2020.

[32] F. Casino, C. Patsakis, E. Batista, O. Postolache, A. Martínez-Ballesté and A. Solanas, "Smart Healthcare in the IoT Era: A Context-Aware Recommendation Example," in *International Symposium in Sensing and Instrumentation in IoT Era (ISSI)*, 2018.

[33] M. I. M. A. F. A. T. T. Erdeniz S.P., "Recommender Systems for IoT Enabled m-Health Applications," in *IFIP Advances in Information and Communication Technology*, 2018.

[34] F. M. M. G. M. e. a. Jabeen, "An IoT based efficient hybrid recommender system for cardiovascular disease," *Peer-to-Peer Networking and Applications,* vol. 12, p. 1263–1276, 2019.

[35] V. S. G. M. ,. R. L. V. V. N. C. ,. D. M. ·. N. Senthilselvan, "An ontology-driven personalized food recommendation in IoT-based healthcare system," *The Journal of Supercomputing,* vol. 75, pp. 3184-3216, 2018.

[36] P. K. ,. D. K. ,. S. R. I. ,. N. U. ,. S.-j. Y. ,. K. K. Farman Ali, "Type-2 Fuzzy OntologyÐaided Recommendation Systems for IoTÐbased Healthcare," *Computer Communications,* 2017.

[37] A. S. Amelie Gyrard, "IAMHAPPY: Towards an IoT knowledge-based cross-domain well-being recommendation system for everyday happiness," *Smart Health Elsevier,* vol. 15, 2020.

[38] P. Z. ,. K. D. M. S. H. M. F. A. Shiqi Yang, "emHealth: Towards Emotion Health Through Depression Prediction and Intelligent Health Recommender System," *Mobile Networks and Applications,* vol. 23, p. 216–226, 2017.

[39] R. M. Frey, R. Xu and A. Ilic, "A Novel Recommender System in IoT," in *5th International Conference on the Internet of Things (IOT 2015)* , Seoul, South Korea, 2015.

[40] S. A. F. A. R. A. Saman Forouzandeh, "Recommender system for Users of Internet of Things (IOT)," *IJCSNS International Journal of Computer Science and Network Security,,* vol. 17, no. 8, 2017.

[41] M. P. R. M. W. L. H. T. H. C. k. M. Sangwhan Cha, "The role of an IoT Platform in the Design of Real-time Recommender Systems," in *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*, 2016.

[42] T. G. Y. S. L. C. ,. L. S. Chenzhong Bin, "A Travel Route Recommendation System Based on Smart Phones and IoT Environment," *Hindawi Wireless Communications and Mobile Computing,* p. 16, 2019.

[43] S.-J. L. Pang-Ming Chu, "A Novel Recommender System for E-Commerce," in *10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI 2017)*, 2017.

[44] F. Lin, Y. Zhou, I. You, J. Lin, X. An and X. Lü, "Content Recommendation Algorithm for Intelligent Navigator in Fog Computing Based IoT Environment," *IEEE Access,* vol. 7, pp. 53677 - 53686, 2019.

[45] S. R. Sergio Di Martino, "An Architecture for a Mobility Recommender System in Smart Cities," *Procedia Computer Science,* vol. 98, p. 425 – 430, 2016.

[46] D. Puiu, S. Bischof, B. Serbanescu, S. Nechifor, J. Parreira and H. Schreiner, "A Public Transportation Journey Planner Enabled by IoT Data Analytics," in *20th Conference on Innovations in Clouds, Internet and Networks (ICIN)*, 2017.

[47] Y. K. Thanh Dinh, "A Novel Location-Centric IoT-Cloud Based On-Street Car Parking Violation Management System in Smart Cities," *Sensors,* 2016.

[48] P. D. H. S. L. P. J. Z. Alireza Hassani, "Querying IoT Services: A Smart Carpark Recommender Use Case," in *IEEE 4th World Forum on Internet of Things (WF-IoT)*, 2018.

[49] V. A. Benjawan Srisura, "Periodical Mobile Recommendation Toward Parking Conflict Reduction," in *IEEE 6th International Conference on Industrial Engineering and Applications*, 2019.

[50] Z. A. a. I. U. Shah Khusro, "Recommender Systems: Issues, Challenges, and Research Opportunities," in *Information Science and Applications(ICISA)*, 2016.

[51] M. Z. X. L. Y. L. S. M. Thiago Silveira, "How good your recommender system is? A survey on evaluations in recommendation," *International Journal of Machine Learning and Cybernetics,* pp. 813-831, 2019.

Dari Speech Classification Using Deep Convolutional Neural Network

Mursal Dawodi
School of Engineering and Science
University of the Ryukyu
Okinawa, Japan

Jawid Ahamd Baktash
Faculty of Computer Science
Kabul University
Kabul, Afghanistan

Tomohisa Wada
School of Engineering and Science
University of the Ryukyu
Okinawa, Japan

Najwa Alam
Faculty of Computer Science
Kabul University
Kabul, Afghanistan

Mohammad Zarif Joya
Faculty of Computer Science
Kabul University
Kabul, Afghanistan

***Abstract*— Recently, speech recognition is one of the most advanced research topics in the world. Many recent research papers have proven the power of deep neural networks in speech recognition systems. The main purpose of this paper is to identify isolated words in Dari speech using deep learning algorithms. This research is one of the new studies in Dari speech recognition and focuses on one-word speech recognition. This collection uses our audio files as a database because there were no Dari language databases on the market at that time. In this paper, the Convolutional Neural Network (CNN) is implemented to detect automatically isolated words in Dari. Besides, it uses Mel frequency coefficients (MFCC) to learn the representation of features during training. This model achieved 88.2% in the test set. The results show that the model can predict samples of words seen during training with high accuracy. However, it is somewhat trying to generalize terms outside the scope of training data and very noisy examples.**

Keywords—speech classification, Dari, convolutional neural network, deep neural network, speech recognition

I. INTRODUCTION

Anyone using technology devices is usually faced with speech recognition systems such as automatic dictation, human-computer dialogue, speech-to-speech translation, system control, etc. Researchers have already evaluated many alternative algorithms to enhance this technology and discover more precise algorithms. In the past few years, deep learning, a specific subfield of machine learning has been a subject of intense media hype and achieved impressive results in speech recognition like its outstanding outcomes in alternative natural language processing themes, image processing, and other research articles [11, 12, 13].

Although there are several studies in speech recognition technology for numerous human languages, there is still a gap for Dari Language. This study develops an automatic speech recognition system in the context of Dari speeches using MFCC for feature extraction and deep CNN for prediction and classification tasks. Moreover, there is not any dataset of Dari language for this context. Therefore, in the early stages of this analysis, we prepared a collection of 2753 audio records.

II. RELATED WORKS

Much research has been done on the implementation of deep neural networks in speech recognition programs. As an example, A. Graves et. Al. [1] represented the efficiency of combining deep, bidirectional Long Short-term Memory (LSTM) RNNs with end-to-end training and weight noise on the TIMIT phoneme recognition benchmark that reduced test set error rate to 17.7%. Similarly, [3] showed that CNN outperforms in speech recognition context and reduces the error-rate. Some of the studies focused on improving accuracy for large vocabulary speech recognition while some other researchers focused on leveraging the power of deep learning in context-based speech recognition. G. E. Dahl et. al. [2] proposed a novel context-dependent pre-trained deep neural networks model. Their investigated model had higher sentence accuracy and lower relative error compared to the former proposed conventional context-dependent Gaussian mixture model.

In recent years some natural processing language researchers started research on Persian NLP. They used different methods and developed several models. One of the best projects is Navisa [8] Persian continuous speech recognition system. [7] proposed a model for isolating Persian spoken words that were evaluated on the first interactive robot with Persian commands in 2008. Another research [4] employed MFCC and a multilayer perceptron feed-forward artificial neural network to distinguish vowel and consonant characters in Persian speeches. They used the Persian Consonant-Vowel Combination dataset that consists of 20 sets of audio samples from 10 speakers and implemented MFCC on every partitioned sound sample. Farsi LTS (letter to the sound system) [5] is a system that translates Persian letters and words to speak. The researchers used neural networks with rule-based and MLP layers. Subsequently, they obtained 61% to 87% conversion accuracy. Likewise, S. Malekzadeh [6] utilized MLP deep learning algorithms to detect Persian phonemes by artificial intelligence and improve voice signal processing.

III. DATA PREPARATION

This research formed a new data set called Isolated Dari Speech Database, the first Dari-speaking audio data set, because we could not find any product on the Internet, as mentioned in the first section. The dataset contains 2753 one-second long utterances of 20 short words, by several people including both male and female. The files are classified into 20 classes each represents a common word or command in Dari. The categories are Gol, Salaam, Qand, Jaan, Nek, Dast, Paa, Bebin, Boro, Beeyaa, Beradar, Khaahar, Haazer, Bad, Bogoo, Boot, Cheshm, khoob, Boo, and Goft. Fig. 1 lists all classes within our dataset and Fig. 2 demonstrations the distribution of each class. In addition to preparing the audio sets, data augmentation was used to reduce overfitting,

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

artificially synthesizing new training data and increase the size of training sets. We structure our data in a csv file where name of audios is listed with their specified label. Fig. 3 shows the metadata for each record of the employed corpus. Similarly, Fig.4 demonstrates an example of depicting words within the corpus as the raw waveform. This study used supervise learning techniques and employed labeled dataset to train the models.

1	گل	Gol
2	سلام	Salaam
3	قند	Qand
4	جان	Jaan
5	نیک	Nek
6	دست	Dast
7	پا	Paa
8	ببین	Bebin
9	برو	Boro
10	بیا	Beeyaa
11	برادر	Beradar
12	خواهر	Khaahar
13	حاضر	Haazer
14	بد	Bad
15	بگو	Bogoo
16	بوت	Boot
17	چشم	Cheshm
18	بو	Boo
19	گفت	Goft
20	خوب	khoob

Fig. 1. Classes within dataset

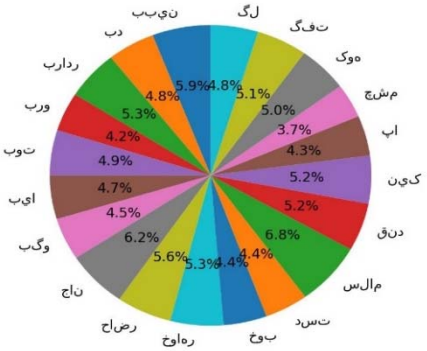


Fig. 2. Class distribution

	slice_file_name	fold	class_name
0	bad1.wav	1	بد
1	bad10.wav	1	بد
2	bad11.wav	1	بد
3	bad12.wav	1	بد
4	bad13.wav	1	بد
5	bad14.wav	1	بد
6	bad15.wav	1	بد
7	bad16.wav	1	بد
8	bad17.wav	1	بد
9	bad18.wav	1	بد

Fig. 3. Metadata example for each prepared Dari sound in the dataset

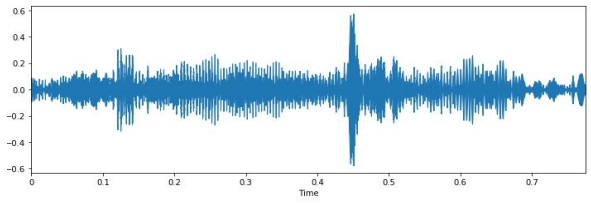


Fig. 4. A sample depicting the word "Jaan" as the raw waveform

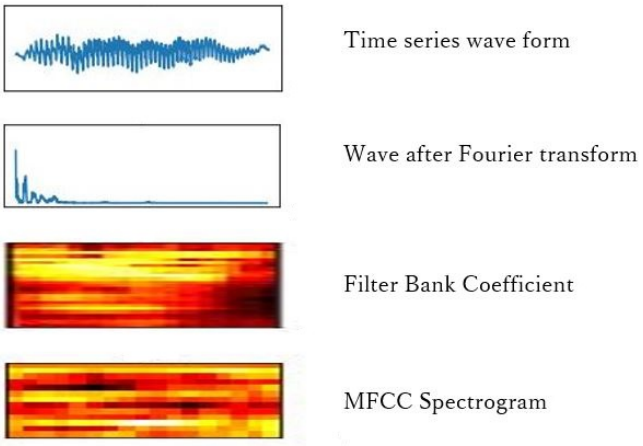


Fig. 5. Time series, Fourier transform, Filter bank coefficient and MFCC of an audio in Paa class

IV. Data Preprocessing

We have pre-processed the audio data before using them. Hence, in the first step, we normalized the sounds to range from -1 to +1. We changed the audio signals to mono because all sounds are considered on one channel only. To extract features from the raw waveform, we used spectroscopy, log-Mel filter banks, and Mel-Frequency implantation coefficients (MFCCs) from stereotyped samples to convert the raw waveform to a time-frequency domain [9] (see Fig. 5). These features are considered as the inputs for neural nets [10]. Moreover, we transformed the data to numeric vectors and subsequently split the dataset into two separate sets of training and validation 80% and 20% respectively.

V. Deep Neural Network Models

Our model is based on CNN to classify utterances and predict associated text. The CNN model consisted of four blocks containing two-dimensional convolutional layers. The following layer is a two-dimensional max pooling layer with a 50% dropout and three dense layers. This model utilizes the ReLU activation function. The number of features, frames, and channels are the input shape of the first layer. The entire layers have a kernel size of 3 throughout the network. The model has 16 convolutional filters in the first layer that is doubled in the proceeding layers. The output layer is a dense layer with an output size of 20 and Softmax activation function as exposed in Fig. 6. The total amount of parameters in the network is roughly 45.892 thousand.

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 39, 173, 16)	80
max_pooling2d_1 (MaxPooling2	(None, 19, 86, 16)	0
dropout_1 (Dropout)	(None, 19, 86, 16)	0
conv2d_2 (Conv2D)	(None, 18, 85, 32)	2080
max_pooling2d_2 (MaxPooling2	(None, 9, 42, 32)	0
dropout_2 (Dropout)	(None, 9, 42, 32)	0
conv2d_3 (Conv2D)	(None, 8, 41, 64)	8256
max_pooling2d_3 (MaxPooling2	(None, 4, 20, 64)	0
dropout_3 (Dropout)	(None, 4, 20, 64)	0
conv2d_4 (Conv2D)	(None, 3, 19, 128)	32896
max_pooling2d_4 (MaxPooling2	(None, 1, 9, 128)	0
dropout_4 (Dropout)	(None, 1, 9, 128)	0
global_average_pooling2d_1 (	(None, 128)	0
dense_1 (Dense)	(None, 20)	2580

Total params: 45,892
Trainable params: 45,892
Non-trainable params: 0

Fig. 6. CNN model architecture summarization

VI. Model Evaluation and Discussion

The model was evaluated based on both training and test set accuracies including 80% and 20% of the samples. This model obtained overall accuracy of 88% and 0.45 validation loss compared as shown in Fig. 7.

For this study, we evaluate the best performance of the models by decreasing and increasing the number of dense layers, pooling layers, dropouts, and kernel size. Finally, the best options were selected for this research as described in previous sections. However, the impressive accuracy result, 88% on testing data, for CNN model was achieved through resampling the audio sample rate to 8000, changing dimensionality of the model to two and increasing the kernel size both in conventional and pooling layers.

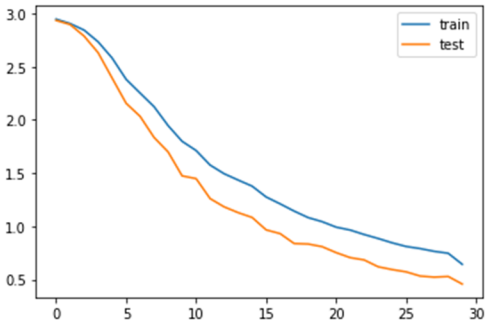


Fig. 7. CNN model: training and evaluation loss

VII. Conclusion and Future Work

In this paper, the effectiveness of the MFCC feature extraction technique and the deep CNN model along with sensible training techniques for Dari one-word Dari speech recognition is demonstrated. This is a preliminary study of Dari natural language processing, and more research is needed to deal with very noisy examples and unsupervised Dari words. In addition, we used the limited vocabulary of Dari words. However, the implementation of more continuous and accurate models requires a very large corpus. Providing a collection of information enriched by the Dari speeches and creating more effective and accurate models are the future work of this research.

References

[1] A. Graves, A. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, May 2013, pp. 6645–6649, doi: 10.1109/ICASSP.2013.6638947.

[2] G. E. Dahl, D. Yu, L. Deng, and A. Acero, "Context-Dependent Pre-Trained Deep Neural Networks for Large-Vocabulary Speech Recognition," IEEE Transactions on Audio, Speech, and Language Processing, vol. 20, no. 1, pp. 30–42, Jan. 2012, doi: 10.1109/TASL.2011.2134090.

[3] O. Abdel-Hamid, A.-R. Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu, "Convolutional neural networks for speech recognition," IEEE/ACM Trans. Audio, Speech and Lang. Proc., vol. 22, no. 10, pp. 1533–1545, Oct. 2014, doi: 10.1109/TASLP.2014.2339736.

[4] S. Malekzadeh, M. H. Gholizadeh, and S. N. Razavi, "Persian Vowel recognition with MFCC and ANN on PCVC speech dataset," arXiv preprint arXiv:1812.06953, 2018.

[5] M. Namnabat and M. M. Homayounpour, "A Letter to Sound System for Farsi Language Using Neural Networks," in 2006 8th international Conference on Signal Processing, Nov. 2006, vol. 1, doi: 10.1109/ICOSP.2006.345518.

[6] S. Malekzadeh, "Phoneme-Based Persian Speech Recognition," arXiv:1901.04699 [cs, eess], Jan. 2019, doi: 10.13140/RG.2.2.32856.96007.

[7] H. Hasanabadi, A. Rowhanimanesh, H. T. Yazdi, and N. Sharif, "A Simple and Robust Persian Speech Recognition System and Its Application to Robotics," in 2008 International Conference on Advanced Computer Theory and Engineering, Dec. 2008, pp. 239–245, doi: 10.1109/ICACTE.2008.125.

[8] H. Sameti, H. Veisi, M. Bahrani, B. Babaali, and K. Hosseinzadeh, "Nevisa, a Persian Continuous Speech Recognition System," in Advances in Computer Science and Engineering, Berlin, Heidelberg, 2009, pp. 485–492, doi: 10.1007/978-3-540-89985-3_60.

[9] Y. Zhang, N. Suda, L. Lai, and V. Chandra, "Hello Edge: Keyword Spotting on Microcontrollers," ArXiv, 2017.

[10] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.

[11] D. Amodei et al., “Deep speech 2: end-to-end speech recognition in English and mandarin,” in Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48, New York, NY, USA, Jun. 2016, pp. 173–182, Accessed: Aug. 22, 2020. [Online].

[12] T. Sainath and C. Parada, “Convolutional neural networks for small-footprint keyword spotting,” 2015.

[13] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” Journal of Machine Learning Research, vol. 15, no. 56, pp. 1929–1958, 2014.

[14] M. Dawodi, T. Wada, and J. Baktash, “An Intelligent Recommender System Supporting Decision-Making on Academic Major,” Information : an international Interdisciplinary journal, vol. 22, no. 3, pp. 241–254, May 2019.

[15] M. Dawodi, T. Wada, and J. A. Baktash, “Applicability of ICT, Data Mining and Machine Learning to Reduce Maternal Mortality and Morbidity: Case Study Afghanistan,” p. 13.

[16] M. Mohammadi, M. Dawodi, W. Tomohisa, and N. Ahmadi, “Comparative study of supervised learning algorithms for student performance prediction,” in 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Feb. 2019, pp. 124–127, doi: 10.1109/ICAIIC.2019.8669085.

[17] M. Dawodi, J. A. Baktash, and T. Wada, “Data-Mining Opportunities in E-Government: Agriculture Sector of Afghanistan,” in 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Oct. 2019, pp. 0477–0481, doi: 10.1109/IEMCON.2019.8936193.

[18] M. Dawodi, M. H. Hedayati, J. A. Baktash, and A. L. Erfan, “Facebook MySQL Performance vs MySQL Performance,” in 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Oct. 2019, pp. 0103–0109, doi: 10.1109/IEMCON.2019.8936259.

Vehicle Type and Color Classification and Detection for Amber and Silver Alert Emergencies Using Machine Learning

Uma K. K. Pillai
Ingram School of Engineering
Texas State University
San Marcos, Texas, USA
u_k9@txstate.edu

Damian Valles
Ingram School of Engineering
Texas State University
San Marcos, Texas, USA
dvalles@txstate.edu

***Abstract*— The National Center for Missing & Exploited Children estimated that 161 AMBER Alerts were issued in the U.S. involving 203 children in 2018, where 85% had involved vehicles and in Florida, 136 Silver Alerts were issued in 2008-2009. The details of broadcasting in Amber and Silver alerts are color, type of the vehicle, vehicle license plate numbers, and car brands. This paper is focused on classifying and detecting vehicle types, colors, and license plates. Currently, a child and older adult were rescued when someone recognized the vehicle in the alert. This paper proposes to design a Machine Learning model to classify the vehicle's colors, types and recognize each vehicle's license plate from camera feeds under different weather conditions and to find possible matches involved in these emergency alerts for the safe return of a child and older adult. Vehicle types include seven classes such as SUV, Sedan, Truck, Bus, Microbus, Minivan, and Motorcycle. Vehicle colors include eight classes: green, blue, black, white, gray, yellow, white, and red. When an Amber or Silver signal is broadcast, the proposed design checks with the vehicle's specifications and extracts the color and type of the vehicle. The model then recognizes the vehicle's license plate of specific vehicle's color and type using image processing techniques and give notification of detected vehicle. Implementing CNN, real-time object detector YOLO, and character recognition model will improve detection and classify vehicle's type, color, and recognize license plate numbers and letters accurately under different environmental conditions for Amber and Silver alert emergencies.**

Keywords*— *Vehicle, classification, detection, CNN, YOLO, Amber alert, Silver alert

I. INTRODUCTION

Various research has been done on vehicle classification using Image Processing and Machine Learning in different applications like traffic monitoring, traffic surveillance systems, and vehicle tracking [1]. The paper is focused on Amber/Silver Alert emergencies. The vehicle's color and type of vehicle are primarily considered for the project. An appropriate color range will need to be defined to process the colored images and match the classification of color for each individual vehicle. The authors in [2] explored different color spaces like the Hue-Saturation-Value (HSV), International Commission on illumination (CIELAB) and Red-Green-Blue (RGB) for vehicle color recognition. These techniques will be considered in this work as an image pre-processing phase that transforms images to improve accuracy performance of the Machine Learning models.

Datasets and pre-processing include those discussed in [2] will be used as it shows challenging due to the noise caused by illumination variation, haze, and overexposure for vehicle's types and colors classification. The first step of the data pre-processing is to apply different image processing techniques to improve image quality. Researchers in [3] and [4] applied dark channel prior (DCP) to remove haze and contrast-limited adaptive histogram equalization (CLAHE) to enhance color contrast. Proper image enhancement and haze removal can help classification systems to obtain better accuracy performance. Features can be extracted using conventional image processing techniques such as Histogram of Oriented Gradients (HOG) [5], color histogram and pre-trained Convolutional Neural Network (CNN).

The proposed approach includes two experimental design methods—first, design two CNN models to classify the vehicle types and another for the vehicle colors. CNN provides proven results in computer vision and image recognition tasks. Second, design a real-time object detector using YOLO to improve detection, and classify vehicle's type and color accurately under different environmental conditions for Amber/Silver alert emergencies. The model, which produces faster detection and classification accuracy, can be integrated within current camera highway systems. The snapshot images will be stored; then, the character recognition model can help identify the numbers and letters in the license plate of specific type and color, and check if it matches the character sequence from the emergency notice. This can help authorities detect target vehicles with a faster detection response and bring a child or elderly safe back home.

II. BACKGROUND

The United States Amber Alert program began in Dallas-Fort Worth, Texas, and it has grown into a program at both state and federal levels [6]. This program was eventually taken to the National Center for Missing & Exploited Children (NCMEC) to request a national initiative. The Amber alert system allows the broadcasters and transportation authorities to immediately distribute information about child abduction to the public through the wireless Emergency Alert System. It also enables the entire community to assist in the search for safe and recovery of the child.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

A Silver Alert is a public notification system in the United States to broadcast information about missing persons – especially senior citizens with Alzheimer's disease, dementia, or other mental disabilities to aid in locating them [7]. The United States Silver Alert program began in Oklahoma, and it has grown into a program at both state and federal levels. A Silver Alert works in much the same way as an Amber Alert for missing children. When a senior or someone with cognitive impairment goes missing, their description and last known whereabouts are sent to law enforcement and media to help bring about their safe return. Public information in a Silver Alert usually consists of the missing person's name and description and a description of the missing person's vehicle. As of December 31st, 2018, there have been 956 successful recoveries of children by the Amber Alert Program, where 85% involved vehicle identification. The retrieval rates resulting from Silver Alert shows more than 90% success level. For example, in North Carolina, 128 Silver Alerts were issued in 2008. Of these, 118 seniors were safely recovered, which results in a 92% success level.

Fig. 1 shows the details of broadcasting in Amber and Silver alerts are color, type of the vehicle, vehicle license plate numbers, and car brands involved in child abduction. Once the alerts are issued, the time taken to recovery is over twelve hours for children in 27 cases. More recovery time is due to the time is taken for someone to recognize the vehicle involved in child abduction and authorities to bring back child or elderly safe back home.

Fig. 1. Broadcasting of an Amber Alert in a Digital Highway Sign [8]

A. Vehicle Colors Classification

The authors in [9] proposed a feature context approach to identify vehicle color. Created their datasets on urban roads, contains 15,601 vehicle images with eight classes of vehicle color. The dataset is very challenging due to the noise caused by illumination variation, haze, and overexposure. Images are taken in frontal view by a high-definition camera. The authors adopted different pre-processing techniques like haze removal to remove haze and color contrast to improve the image's quality. They achieved an average accuracy of 90.68% using the Support Vector Machine (SVM) classifier.

The authors in [10] used the dataset in [9] to classify eight classes of vehicle color using the Convolution Neural Network (CNN). The CNN model learned classification based on color distribution. They convert the input image into two different color spaces using the Hue-Saturation-Value (HSV) and the International Commission on illumination (CIELAB) transformations. They experimented with different color spaces such as Red-Green-Blue (RGB), HSV, and CIELAB. Achieved an accuracy average of 94.47% in RGB color space, which outperforms the dataset in [9] using a feature context approach.

Researchers in [11] used a CNN architecture to classify vehicle types with four classes (small, medium, large, and unknown) and vehicle color with seven classes (black, blue, white, green, yellow, red, and unknown). They extracted 914 vehicle images from surveillance videos. The authors compared their method with previous work that utilized decision trees, random forest, and DNN classifier. The results show that the classification of vehicle type accuracy increased by 1.8%, and vehicle color increased by 0.8%. The drawback is that they did not consider the environmental condition and fine-tuning with different hyperparameters.

B. Vehicle Types Classification

Researchers in [12] proposed a modified YOLOv2 network structure to classify vehicle types with eight classes by a multi-layer feature fusion strategy. The strategy means global and local features of one layer are fused so that model can distinguish tiny differences among vehicle types and removed repeated convolution layers in the higher ends. The authors used two datasets such as the Beijing Institute of Technology (BIT) Vehicle dataset containing 9,580 vehicles, and the Comp-cars Dataset containing commercial vehicle images (1,687 vehicles) and from road surveillance cameras. It contains only Sedan and SUV type of vehicles with just over 40,000 image samples. It includes day and night scenes, but there is no noise background, rain, and other weather conditions. The authors compared their model with YOLOv2 and showed a mean average precision of 94.8%.

The authors in [13] proposed an improved YOLOv2 network structure for Multi-scale detection in varying vehicle sizes. They proposed a Ratio k-means (R k-means++) clustering algorithm that generates anchor boxes to improve location accuracy. They proposed a Focal Loss to reduce the negative influence on training resulting from an imbalance between vehicles and background for training the model to decrease the imbalance between vehicles and background noise by adjusting hyper tuning focusing parameters such as gamma and alpha. Experimented with the BIT-Vehicle dataset and achieved a 97.3% mean average precision. The main drawback is that they did not consider environmental conditions.

C. Licence Plate Recognition

The authors in [14] proposed an automatic number plate recognition system using image processing technology to identify the vehicles. They used various image processing techniques such as Grayscale Conversion, Noise Filtering, Image Binarization, Histogram Equalization, and Template Matching. They recognized characters and numbers in the license plate using MATLAB. Researchers in [15] proposed an Adaptive Technique for Computer Vision-Based Vehicles License Plate Detection System. Three module includes Detection of a license plate, Segmentation of Characters, and Text Box Generation. They achieved an average accuracy of 78.2% using MATLAB.

Deep Learning has proven to be fast technique for vehicle detection and accurately classify vehicle's color and type. The character recognition model will help identify the right vehicle by correlating the license plate of specific vehicle's type and color involved in alerts.

III. Design Process and Cosideration

The workflow model shows in Fig. 2 includes collecting datasets, data pre-processing, feature extraction, vehicle detection, classification, image processing and character recognition.

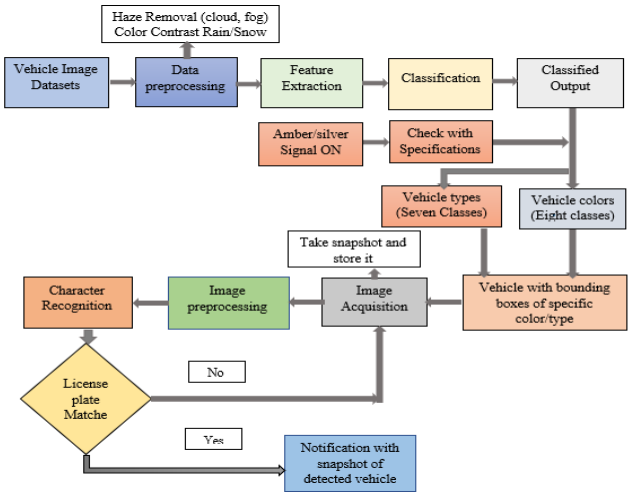


Fig. 2. Workflow of proposed vehicle colors, types classification and License plate recognition Model

A. Datasets

The first step for the Machine Learning model is to collect the best vehicle image datasets considering different environmental conditions like lighting and illumination with different viewpoints. Several datasets will be used for training for vehicle color and vehicle type classification. Some of the datasets are discussed in the following subsections.

a) Comprehensive Cars (CompCars) Dataset: CompCars dataset will be used to evaluate the vehicle type model. It contains some surveillance of 50,000 car images that suffered from significant variations in illumination due to differences in traffic imaging conditions, making the recognition of vehicles from the frontal-view surveillance data more challenging. The actual size of the original images is 800x850 pixels in resolution. It includes twelve types of cars: MPV, SUV, hatchback, sedan, minibus, fastback, estate, pickup, sports, crossover, convertible, and hardtop convertible [16]. Dataset is shown in Fig. 3 can be used for vehicle types and colors by labeling using tools such as LabelImg and Labelbox.

Fig. 3. Examples of surveillance data images from [16]

b) Vehicle Color Recognition Dataset: The authors in [9] used a dataset consisting of 15,601 vehicle images with eight classes of vehicle color, which are black, blue, cyan, gray, green, red, white, and yellow. The dataset is very challenging due to the noise caused by illumination variation, haze, and overexposure. Images are taken in frontal view by a high-definition camera, and Fig. 4 shows the vehicle color classification example. Vehicle color images will be label using labeling tools.

Fig. 4. Examples of images from [9] are noisy and brightness constancy

B. Data Pre-processing

Data pre-processing aims to improve vehicle images by removing haze and enhancing color contrast of image features, thereby improving the performance of the proposed classification. The pre-processed image is fed to the Machine Learning models for training and testing images.

a) Haze Removal: Hazy weather conditions pose the most common problem in surveillance systems, reducing the clarity in traffic images resulting in low accuracy of vehicle recognition and detection. The dark channel prior (DCP) is based on the property of "dark pixels," which have a very low intensity in at least one-color channel, except for the sky region. As for effectiveness in dehazing, the majority of new dehazing techniques have adopted the DCP. The DCP-based dehazing techniques are composed of four significant steps: atmospheric light estimation, transmission map estimation, transmission map refinement, and image reconstruction.

b) Enhance Contrast: It is an enhancement technique to improve the contrast in an image by stretching the range of intensity to obtain the full range of pixel values. Contrast-limited adaptive histogram equalization method CLAHE has proven the highest accuracy of object detection. It will be used for image enhancement. Proper image enhancement can help classification systems to obtain better accuracy performance using OpenCV in Python.

c) Noise Filtering: The image is filtered using median filter to remove the noise, while preserving the sharpness of the image. The filter can be used a non-linear filter where it replaces each pixel with a value obtained by computing the median of values of pixels..

C. Proposed Experiment Method-1

a) Feature Extraction: Feature extraction of vehicle types includes a bounding box (x-coordinate from top left corner, y-coordinate from top left corner, width, and height) and the ratio of a vehicle over background. Feature extraction of color includes different color spaces like HSV and RGB. Features can be extracted using image processing techniques such as HOG, color histogram, and pre-trained CNN. Convolution layers in CNN architecture will be used to extract feature maps. These extracted features will be sent to the classification model and measure accuracy performance using OpenCV in Python.

b) CNN Architecture: The development of the architecture includes defining the number of nodes and size of filters in the convolution layer, pooling layers, activation functions, number of nodes in the fully-connected layers, and number of dropout layers. Two CNN models will be used to classify the vehicle type and another for the vehicle color. The last layer of vehicle type and color classification contains ten neurons each. Designing a proper CNN model will be based on trying several kinds of research approaches applied different CNN models [17], such as the VGGNet and the AlexNet models. The model performance can be increased by tuning the hyperparameters, such as the learning rate, batch size, optimizer, and the number of epochs. Searching CNN's suitable hyperparameters that could make the performance of the classifications be more accurate and stable. Fig. 5 shows the workflow of Experiment Method-1 using CNN.

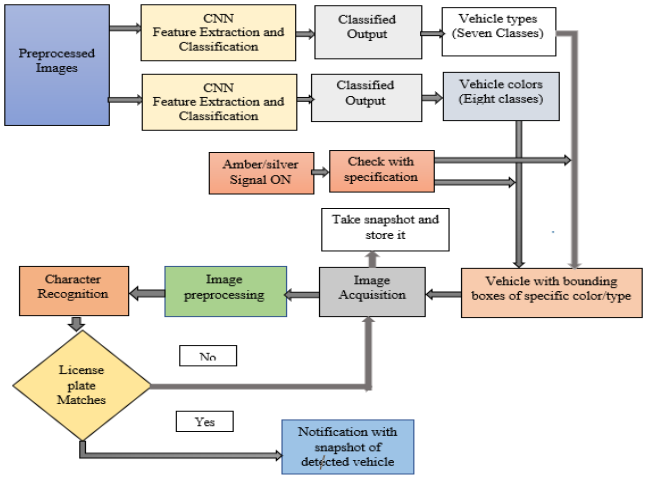


Fig. 5. Workflow of Experiment Method-1 using Two CNNs

c) Character Recognition: The main goal of the proposed Character Recognition model is to recognize the vehicle's license plate letters and numbers accurately from the captured vehicle image. The process of executing the character recognition model includes Image Acquisition and Image Processing. The process implementation is discussed in the following sections.

1) Image Acquisition: The snapshot of the captured image will be used as input to the proposed recognition model.

2) Image Processing: Researchers in [18] applied different image processing techniques to recognize characters accurately. It includes Grayscale Conversion, Image Binarization, Edge Detection, and Template matching. These image processing techniques will be implemented using Python and OpenCV, which are discussed in the following sections.

a) Grayscale Conversion: The first step of processing is to convert the captured image of RGB to grayscale images to identify the characters of the plate easier and make further processing simpler. The gray color is one in which the red, green, and blue components all have equal intensity in RGB space. Therefore, it carries only intensity information for each pixel.

b) Image Binarization: Image binarization is a process to convert an image to black and white. In this method, a certain threshold is chosen to classify specific pixels as black and individual pixels as white. Thresholding is to segment the license plate characters from the background.

c) Edge Detection: Edge location is the principal technique for feature extraction. Distinctive edge identification such as Canny, Sobel, Prewitt, and Roberts cross can be utilized for edge recognition. Edge detection is to discover pixels in the image where edges are searching for discontinuities in angles.

d) Template Matching: The Template Matching technique is to recognize characters accurately and shows the similarities between a given template and the image that should be matched with it. The character recognition model can identify numbers and letters in the license plate of specific type and color, and check if it matches with the character sequence from the emergency notice.

D. Proposed Experiment Method-2

In the proposed experiment Method-2 YOLO9000 will be used for vehicle color and type classification. Fig. 6 shows the workflow of Experiment Method-2 using YOLO9000. The rest of the process is described back on the Character Recognition subsection C from Method-1.

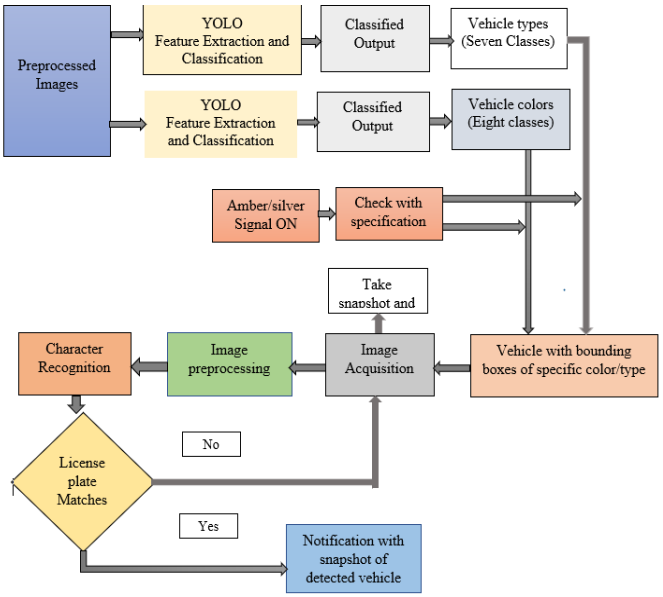


Fig. 6 Workflow of Experiment Method-2 using YOLO9000

Researches in [19] proves faster detection on different dataset like PASCAL Visual Object Classes (VOC 2007) and Common Objects in Context (COCO) when compared to other detectors such as SSD and R-CNN. YOLOv2 bounds the location using logistic activation σ, which makes the value fall between zero to one. The network predicts five bounding boxes at each cell in the output feature map. The network predicts five coordinates for each bounding box, t_x,t_y,t_w,t_h, and t_o. If the cell is offset from the top left corner of the image by (c_x, c_y) and the bounding box prior has width and height (p_w, p_h), then the predictions are calculated with (1) and (2), where, IOU is

the overlap rate between the bounding box detected by the system and the ground truth box. BB_{gt} is the ground truth box on training label. BB_{dt} is the detection bounding box as shown in Fig. 7.

$$p_r(object) * IOU(b, object) = \sigma(t_o) \quad (1)$$

$$IOU = \frac{area(BB_{dt} \cap BB_{gt})}{area(BB_{dt} \cup BB_{gt})} \quad (2)$$

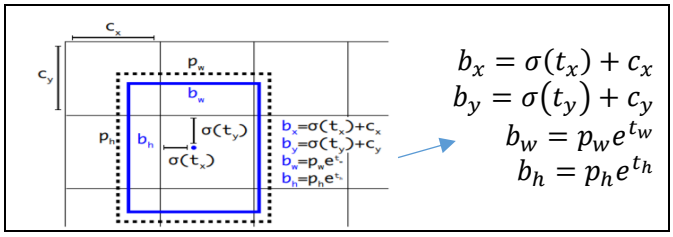


Fig. 7. Bounding Box Prediction from [19]

Testing will be carried out in two ways. First, the model will be tested with traffic surveillance videos [20] to identify the number of frames taken to place bounding boxes around vehicles and accurately recognize specific colors and types. Second, the model will be tested running with an NVIDIA Jetson Nano Developer Kit [21]. Fig. 8 shows the Jetson Nano with several USB ports that will serve the implementation.

Fig. 8. NVIDIA Jetson Nano [21]

IV. Conclusion

Accomplishing high accuracy of vehicle type and colors when dealing with Amber and Silver Alerts under different weather conditions using Machine and Deep Learning techniques. The proposed methods-CNN, YOLO, and Character Recognition model will be implemented to improve classification accuracy, faster detection, and recognize license plates accurately. This can help authorities to detect target vehicles with a faster detection response and bring a child or elderly safe back home.

References

[1] Jun-Wei Hsieh, Shih-Hao Yu, Yung-Sheng Chen and Wen-Fong Hu, "Automatic traffic surveillance system for vehicle tracking and classification," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 2, pp. 175-187, June 2006.

[2] C. Hu, X. Bai, L. Qi, P. Chen, G. Xue and L. Mei, "Vehicle Color Recognition With Spatial Pyramid Deep Learning," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2925-2934, Oct. 2015.

[3] J. Hu, Z. Li and X. Chen, "Modified Image Haze Removal Algorithm Based on Dark Channel Prior," *2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*, Xiamen, China, 2019, pp. 1600-1605.

[4] D. S. Chandana, K. Chigurupati, A. Srikrishna and B. Venkateswarlu, "An Optimal Image Dehazing Technique Using Dark Channel Prior," *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, Kannur,Kerala, India, 2019, pp. 609-614.

[5] N. Habibunnisha, K. Sivamani, R. Seetharaman and D. Nedumaran, "Reduction of Noises From Degraded Document Images Using Image Enhancement Techniques," *2019 Third International Conference on Inventive Systems and Control (ICISC)*, Coimbatore, India, 2019, pp. 522-525.

[6] "National Center for Missing & Exploited Children," *Wikipedia.*Accessed April2,2020.[Online].Available:https://en.wikipedia.org/wiki/National_Center_for_Missing_%26_Exploited_Children.

[7] "Silver Alert," *Wikipedia.* Accessed April 2, 2020. [Online]. Available: https://en.wikipedia.org/wiki/Silver_Alert.

[8] "Amber Alert,"*Wikimedia*. Accessed April 15, 2020. [Online]. Available: https://commons.wikimedia.org/wiki/File:Amber_Alert.jpg.

[9] P. Chen, X. Bai and W. Liu, "Vehicle Color Recognition on Urban Road by Feature Context," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2340-2346, Oct. 2014.

[10] Rachmadi, R. and Purnama.I,"Vehicle Color Recognition using Convolutional Neural Network," arXiv, 2015. [Online]. Available: https://arxiv.org/format/1510.07391.

[11] W. Maungmai and C. Nuthong, "Vehicle Classification with Deep Learning,"*2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*, Singapore, 2019, pp. 294-298.

[12] Sang, J., Wu, Z., Guo, P., et al., "An Improved YOLOv2 for Vehicle Detection," *Sensors (Basel)*. 2018.

[13] Wu, Zhongyuan, Jun, Qian, Hong, Bin and Xiaofeng, "Multi-Scale Vehicle Detection for Foreground-Background Class Imbalance with Improved YOLOv2" *Sensors*, 2019.

[14] P.Meghana, S. SagarImambi, P. Sivateja and K. Sairam "Image Recognition for Automatic Number Plate Surveillance," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Volume-8 Issue-4, February 2019.

[15] D. Chowdhury, S. Mandal, D. Das, S. Banerjee, S. Shome and D. Choudhary, "An Adaptive Technique for Computer Vision Based Vehicles License Plate Detection System," *2019 International Conference on Opto-Electronics and Applied Optics (Optronix)*, Kolkata, India, 2019, pp. 1-6.

[16] L. Yang, P. Luo, C. C. Loy and X. Tang, "A large-scale car dataset for fine-grained categorization and verification," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 2015, pp. 3973-3981.

[17] Mingyang Zhang, Pengli Wang, and Xiaoman Zhang, "Vehicle Color Recognition Using Deep Convolutional Neural Networks,"*International Conference on Artificial Intelligence and Computer Science*, July 2019. [Online]. Available: https://dl.acm.org/doi/10.1145/3349341.3349408

[18] P.Meghana, S. SagarImambi, P. Sivateja and K. Sairam "Image Recognition for Automatic Number Plate Surveillance," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Volume-8 Issue-4, February 2019.

[19] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017, pp. 6517-6525, doi: 10.1109/CVPR.2017.690.

[20] D. Zapletal and A. Herout, "Vehicle Re-identification for Automatic Video Traffic Surveillance," in *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Las Vegas, NV, pp. 1568-1574, 2016.

[21] "Jetson Nano Developer Kit," Accessed April 2, 2020. [Online]. Available:https://developer.nvidia.com/embedded/jetson-nano-developer-kit.

Detecting Physical Activities Using Body-Worn Accelerometers

Sudarshan S. Chawathe
School of Computing and Information Science & Climate Change Institute
University of Maine
Orono, Maine 04469-5711, USA
chaw@eip10.org

***Abstract*—This paper addresses the task of using data from accelerometers attached to a person's body to determine the kind of physical activity being performed by that person. The activities of interest are routine ones such as sitting, walking up a flight of stairs, walking, and jogging. The paper describes methods for segmenting the time-series data from accelerometers and for extracting features that are effective for determining activities when used in conjunction with well established classification algorithms. These methods are implemented in a prototype that is used to evaluate their effectiveness on a publicly available dataset of tagged accelerometer traces. The prototype also provides intuitive visualizations of the accelerometer traces, allowing a human expert to gain a better understanding of both the dataset and the predictions from the classifiers. Although the methods in this paper use fewer and simpler features extracted from the raw accelerometer data, they provide higher accuracies when compared to those reported in prior work on the experimental dataset.**

***Index Terms*—Human Physical Movement; Activities of Daily Living; Accelerometers; Classification.**

I. Introduction

Automated and continual detection of the physical activity being performed by an individual is often important in healthcare. For instance, such detection may be used to determine the amount of exercise that an individual is getting or to detect unhealthy patterns of exertion or physical stress after a medical procedure. Another important application is determining the extent to which an individual is engaging in so-called activities of daily living, which have important implications for health and well-being [1]–[3].

There is a wide variety of methods that may be used for detecting physical activities of individuals. These include methods that are only partly automated and rely on a human for some part of the detection, such as remote video monitoring. Such methods may be suitable when there are other reasons for remote monitoring by a human. However, in many situations, the effort and cost they entail is prohibitive. Therefore, fully automated methods for detecting physical activities are of particular importance.

Methods for automated detection of physical activities may be categorized based on the kinds of sensing technologies and devices they use, and the manner in which sensors are deployed. For instance, simple, low-cost binary sensors (such as contact switches) that are installed on a a few doors and thresholds in a person's home can often provide sufficient information for determining the activity of daily living in which that person is engaged [4]. Apart from low cost and ease of deployment, a significant benefit of these sensors is that they are unobtrusive and require no physical contact with the individual. However, these sensors do not provide sufficient information for purposes such as monitoring exercise levels due to their limited temporal granularity in particular. On the other hand, by attaching a variety of intrusive sensors to a person's body, it is possible to get extremely detailed information on variables such as heart rate, breathing rhythm, and blood oxygen saturation levels. In this context, accelerometers provide an attractive design point as they are inexpensive and only modestly intrusive and can provide data at a fine temporal granularity.

Given the prevalence of smartphones that include accelerometers, there has been work on using those devices for the purposes of detecting activities, in particular falls or other problematic situations [5]. While the accuracy provided by those devices and, in particular, their deployment (manner in which they are carried by individuals) may be sufficient for some applications, other applications demand the higher reliability and accuracy that may be achieved by special-purpose accelerometers that are attached at specific locations on the body. This paper focuses on the use of such specially deployed accelerometers for detecting physical activities.

The *main question addressed by this paper* is: What are effective methods of converting the raw data from accelerometers attached to an individual's body into predictions of the physical activity in which that individual is engaged?

The main *contributions* of this paper may be summarized as follows:

- It describes the problem of mapping raw accelerometer data to predictions of physical activity.
- It provides a method for extracting a modest number of easily computed features from raw accelerometer data.
- It summarizes the results of an experimental evaluation of the effectiveness and efficiency of methods that use these features to predict physical activity.
- It outlines the design and implementation of a prototype system for interactively exploring accelerometer traces.

The main *results* may be summarized as follows:

978-1-7281-9615-2/20/$31.00

- Although the set of features used by this work is simpler than that in some earlier work, the resulting predictions provide higher accuracies (cf. Fig. 3 and [6, Fig. 8]).
- The computational costs of these methods are very modest, permitting near real-time detection of activities using low-cost commodity hardware that may be easily deployed.

Paper outline: The task of predicting physical activities from accelerometer data is detailed in Section II. It describes the characteristics of the input data using concrete examples drawn from a publicly available dataset from prior work [6] and ends with a problem definition in this context. Section III addresses the task of transforming the raw time-series data from accelerometers into a form that is more amenable to classification methods for use in predictions. In particular, it describes how the continuous time-series data is segmented into instances and defines the set of features (attributes) extracted from the time series segment for each instance. Section IV describes how these features are used in conjunction with well established classification algorithms to predict physical activities and summarizes some results from an experimental study using the primary dataset. In order to augment the quantitative insights of this section with qualitative and intuitive observations, Section V uses a few illustrative examples of visualizations of accelerometer traces provided by the system prototype. Related work is discussed in Section VI. Section VII summarizes this work and briefly notes topics of ongoing and future work.

II. Activities from Accelerometer Traces

For concreteness, the following focuses on a setup composed of two accelerometers attached to each individual: one on the wrist and one on the thigh, as used to generate the dataset from prior work [6] that is used in the experimental study. Each accelerometer is sampled at 100 Hz in each of the three spatial axes, yielding six traces.

Fig. 1 depicts a visualization of such traces, produced by the prototype system described later. Top to bottom, the first three plots are the x, y, and z accelerations for the accelerometer attached to the thigh and the next three plots are the similar accelerations for the accelerometer attached to the wrist. The horizontal axis marks time in units of 10 seconds per tick mark and the vertical axes for each plot denote acceleration. The larger tick mark to the left of each plot marks the origin (0 m/s^2) for that plot and the smaller mark next to it marks 0.2 m/s^2. The scale of the vertical axis for each of the six plots varies and may be inferred by observing the distance between the larger and smaller tick marks to its left. For example, we may note that the tick marks on the vertical axis adjacent to the third plot (cyan, for the z axis of the thigh accelerometer) are spaced much closer than the similar tick marks adjacent to the sixth plot (black, for the z axis of the wrist accelerometer), indicating that, during the plotted period, the z accelerations recorded by the thigh accelerometer have a larger amplitude than those recorded by the wrist accelerometer.

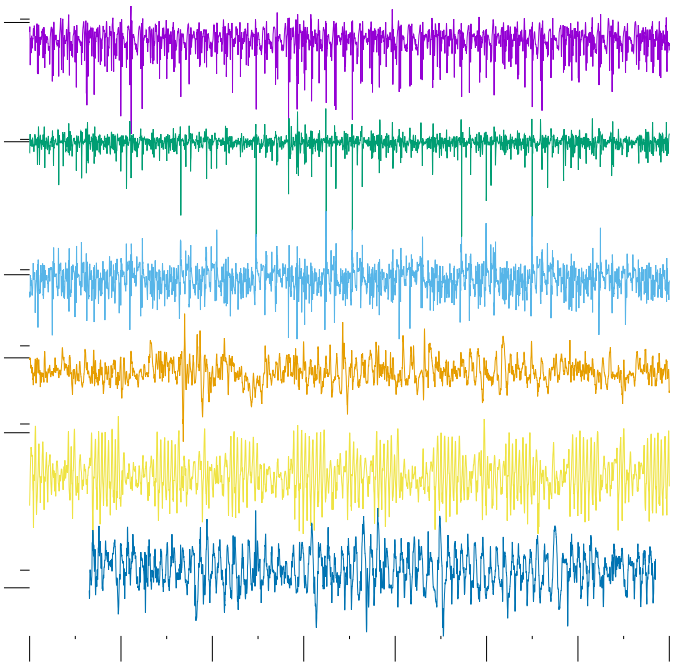


Fig. 1. Accelerometer traces, top to bottom: x, y, and z accelerations for the thigh sensor, followed by x, y, and z accelerations for the wrist sensor. The scale on the horizontal axis is 10 seconds per (larger) tick mark. The scale and origin on the vertical axis is different for each of the six plots. The vertical origin (acceleration of 0 m/s^2) for each plot is marked by the larger tick mark to its left. The scale may be inferred from distance between that larger tick mark and the smaller one adjacent to it, which marks 0.2 m/s^2.

The dataset due to prior work [6] includes such accelerometer traces generated by 33 participants performing each of six physical activities for approximately a minute each. (The participants are identified only by a numeric code and no additional information, such as age or gender, is available.) The activities are summarized in Table I. Each sample in the dataset is a six-tuple of values for the x, y, and z accelerations of the thigh and wrist accelerometers, nominally at 100 Hz, along with a timestamp, and is tagged with the activity that the participant was performing at that time.

TABLE I
Physical activities performed by each of 33 participants.

act. ID	activity
0	walking downstairs
1	jogging
2	lying
3	sitting
4	standing
5	walking upstairs

In this context, the task of detecting physical activities consists of predicting the activity at each sampled instant of time using only the accelerometer readings.

III. From Accelerometer Traces to Instances and Features

Following common practice for using time series data to perform predictions, such as the prediction described at the end of the previous section, we transform the time series data into *instances*, which are the objects that will be used for both training prediction models and for the predictions themselves. Such transformation is motivated by the observation that an individual sample (six-tuple of x, y, z accelerations for the thigh and wrist accelerometers) is best interpreted in relation to other temporally nearby samples. This observation also motivates the use of a temporal window to define instances. The width (duration) of the window is an important design decision. Longer windows provide more contextual information but incur a higher computational overhead and cause a prediction delay equal to at least the window length. (The delay may be reduced by the use of partly overlapping staggered windows, but at the expense of an even larger computational cost because each sample from the original trace appears in multiple instances instead of just one.) Smaller windows have lower computational costs and enable shorter response times but may miss some contextual information from samples that are farther apart. This paper defines instances using 10-second non-overlapping (abutting) windows, following suggestions from prior work [6].

TABLE II
Features computed per 10-second window (non-overlapping, abutting) for the six-fold accelerometer traces (x, y, z accelerations for the thigh and wrist sensors). The complete name of each feature is obtained by prefixing the abbreviation in the first column of the table with two letters: the first (t or w) identifies the sensor (thigh or wrist) while the second (x, y, or z) identifies the spatial axis. Thus, for instance, txac05 denotes the shift-by-5 autocorrelation feature for the x axis acceleration trace from the thigh accelerometer.

abbr.	feature	count
n	number of samples	1
mn	mean value	1
va	sample variance	1
sk	skewness	1
ku	kurtosis	1
ac_i	autocorrelation with shift $i = 1, \ldots, 11$	11
ft_i	Fourier Transform term $i = 1, \ldots, 11$	11
	total:	27

For each instance determined by the 10-second sampling window, each sensor, and each axis, the 27 features described by Table II are computed, for a total of $2 \times 3 \times 27$ features per instance. In this way, the prediction task described at the end of the previous section is transformed into a classification problem that may be addressed using well established classification algorithms such as Decision Tables and Random Forest, as is described below.

IV. Experimental Study

We may use a variety of algorithms to perform the classification task described at the end of the previous section. This section summarizes the results of an experimental study of such classification using the algorithms listed in Table III.

TABLE III
Classification algorithms used in experimental study

abbr.	classifier
ZerR	0-R: ignores attributes and selects most frequent class
OneR	1-R: predicts using a single attribute selected during training
DTab	Decision Table
PART	PART Decision List
JRip	RIPPER [7]
J48	C4.5 decision tree
RTre	Random Tree
RFor	Random Forest
NBay	Naive Bayes
BNt1	Bayes Net with single parent per node
BNt5	Bayes Net with up to five parents per node
IBk	Instance Based classifier: k-nearest neighbor

In order to enable experimental study, the methods described earlier were implemented on the JVM (Java Virtual Machine) platform using the Kawa Scheme programming environment [8]. The experiments were performed on a notebook computer with an Intel Core i5-5250U CPU (2 cores, 4 hyper-threads), 8 GiB of RAM, and a SSD SM0256G for external storage, running the Debian 9 (stretch) GNU/Linux distribution and OpenJDK 1.8. The WEKA workbench's implementation (v. 3.8.2) of the conventional classification algorithms of Table III is used.

TABLE IV
Confusion matrix for the classifier of Fig. 2.

activity	activity (by two-letter prefix)					
	st	ly	si	up	do	jo
*st*anding	634	0	0	3	1	0
*ly*ing	0	611	15	0	0	0
*si*tting	0	15	598	0	0	0
*up*stairs	0	0	1	262	30	4
*do*wnstairs	0	0	0	19	247	6
*jo*gging	1	0	0	4	7	805

The results of using these algorithms with the instances and features defined in the previous section to predict physical activities are summarized by Figs. 3 to 8. For these experiments, both the participant IDs and timestamp attributes were removed from the dataset. Without such removal, the activity can be predicted with high accuracy based on the participant and timestamp. While technically valid, such a prediction is not usable in practice for the envisioned applications where the times at which a person performs various activities are not known in advance.

As summarized by Figs. 3 to 5, several classification algorithms provide very high accuracies, indicating the effectiveness of the features defined earlier. In particular, these accuracies are higher than those reported for similar experiments in in prior work [6, Fig. 8] despite the latter using a more sophisticated set of features. Figs. 6 to 8 summarize the performance aspects of these classifiers. All three charts

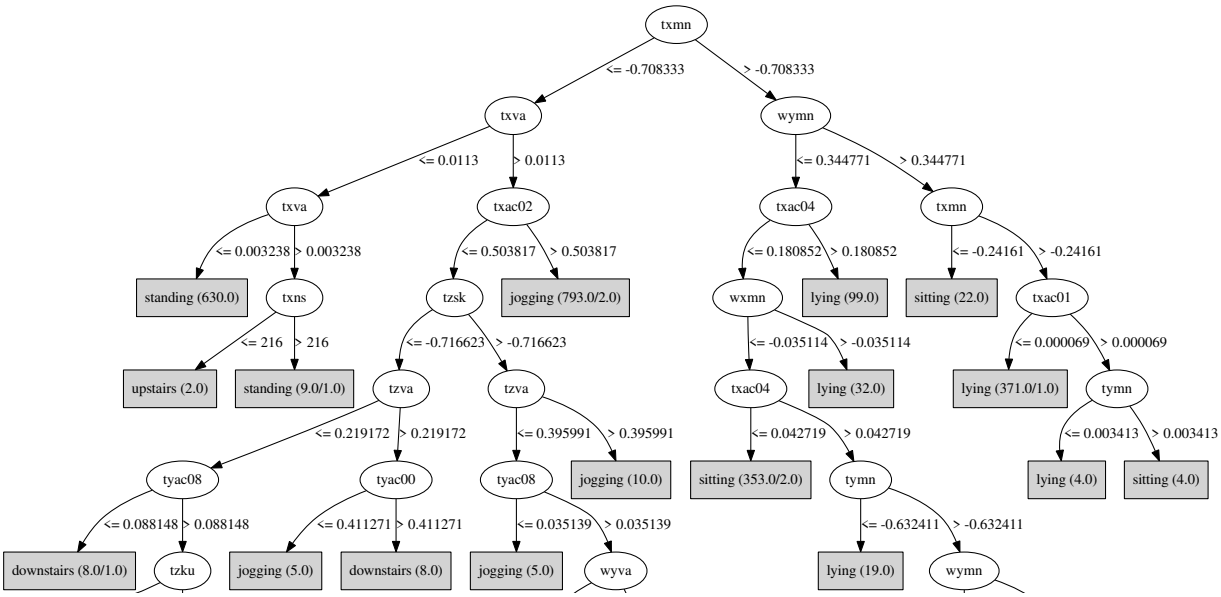


Fig. 2. Tree classifier generated by J48, truncated at the bottom for brevity. The depth of the full tree is 12. The classifier accuracy is 96.75%.

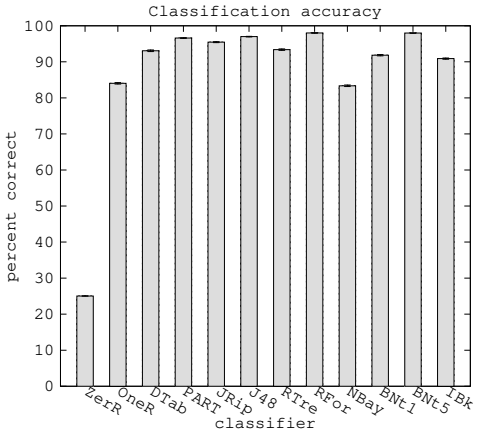


Fig. 3. Percentage of instances correctly classified by method. The error bars (barely perceptible) in this and the following charts denote standard error of the mean.

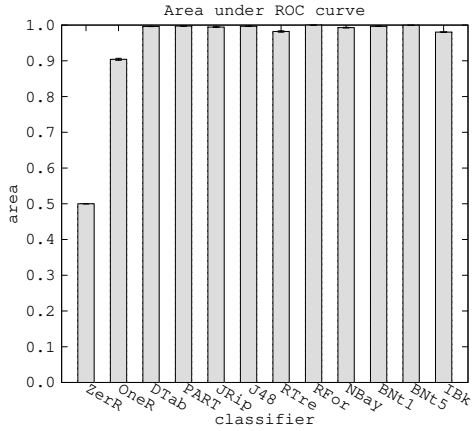


Fig. 5. Area under the ROC (receiver operating characteristic) curve by method.

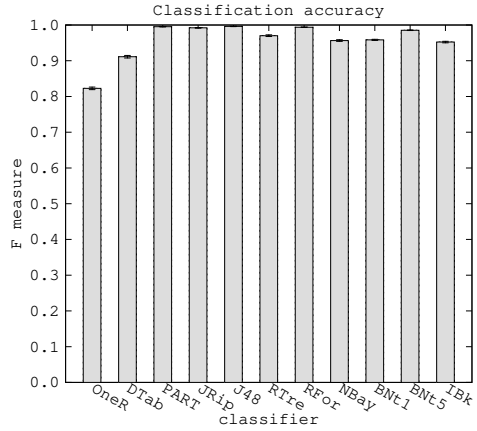


Fig. 4. Classification F-measure by method.

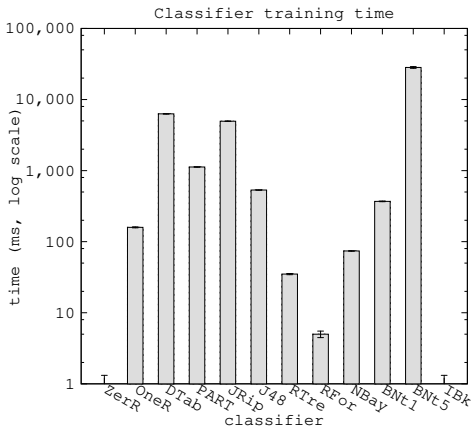


Fig. 6. Training (model building) time by method.

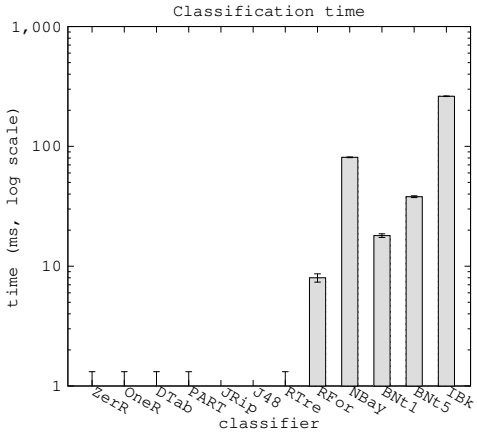


Fig. 7. Testing (classification) time by method.

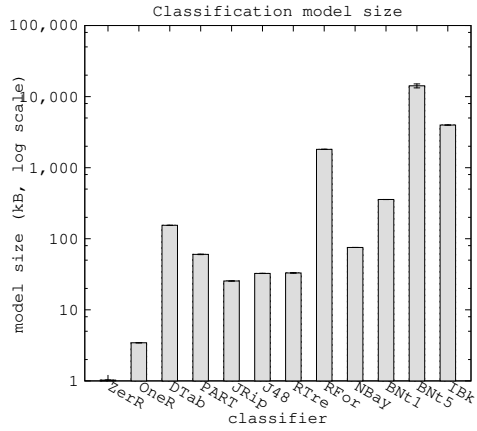


Fig. 8. Learned model size by method.

have a logarithmic scale on the vertical axis. Although there is a large variation in training times (Fig. 6), the overall times are sufficiently low to not be of practical concern. The testing (prediction) times summarized by Fig. 7 are too small to plot meaningfully except for a few methods, and are well under a second for even the slowest method. The sizes of the learned models are summarized by Fig. 8.

A notable advantage of tree (non-ensemble) and rule-based classifiers is their amenability to human comprehension, which permits a human to validate the classifier and to investigate and fine-tune thresholds and other aspects. Fig. 2 depicts a portion of the C4.5-style [9] classification tree generated by the J48 implementation that is part of the WEKA workbench [10], yielding a classifier with high accuracy on the dataset under study. The confusion matrix for this classifier appears in Table IV. Another method that both has good prediction accuracy and is easy for humans to understand is PART, and an excerpt of that classifier appears in Fig. 9.

V. Visualizing Traces

In order to gain a better understanding of the quantitative results summarized by the charts in the previous section, as well as to facilitate exploratory work on data from accelerometers,

```
txmn > -0.708333 AND
wymn > 0.344771 AND
txmn > -0.24161 AND
txac01 <= 0.000069: lying (371.0/1.0)

txmn > -0.708333 AND
txac04 <= 0.180852 AND
wxmn <= -0.025767 AND
txac04 <= 0.042719 AND
wxmn <= -0.181665 AND
wzva <= 0.022003: sitting (358.0)

txmn > -0.708333 AND
tzmn > -0.837136: lying (122.0)

txmn > -0.708333 AND
wymn <= -0.138346 AND
wzva <= 0.033749: lying (63.0)

tzmn <= -0.643647 AND
txft00 > -1214.109375 AND
wzmn > 0.138142 AND
wxmn <= -0.700811: sitting (164.0)
```

Fig. 9. Classification rules generated by PART on the dataset with both participant ID and time attributes removed, truncated to the first five rules for brevity. There are 21 rules in all and the classification accuracy is 96.87%.

part of the work leading to this paper has developed a prototype system called *ATV* (for Accelerometer Traces Visualizer). The non-interactive components of the same system are also used for the experiments reported in the previous section. This system for visualizing traces enables quantitative results to be qualitatively validated by providing interactive comparison of traces in several ways.

For instance, Fig. 10 depicts a screenshot of the system displaying the accelerometer traces on a per-participant basis. The six charts displayed (one for each activity) per participant enable, for example, the intuitive determination of the levels of similarities and differences across activities. The charts use the same conventions as those used by the chart of Fig. 1.

In contrast, Fig. 11 depicts a screenshot of the system being used to compare traces across participants for a single activity, in order to gain an intuitive understanding of the accelerometer signature of that activity.

The implementation uses a fully standards-based Web user interface. Data and layout are cleanly separated using XHTML and CSS. The charts rendered in SVG (scalable vector graphics) format. Much of the interactivity is achieved using precomputed static Web pages that are suitably linked. The remainder can be achieved using client-side JavaScript and related technologies. Features such as zooming in and out of portions of the rendered traces can be implemented completely on the client side, not only enabling very responsive interactions but also eliminating load on the back-end system. The back-end processing is performed on the JVM platform using the Kawa Scheme programming environment [8]. Historical and new data is stored in a SQLite database [11].

VI. Related Work

There are diverse kinds of time-series data generated by human activities and a significant body of work dedicated to their analysis. For instance, in the audio domain, there has

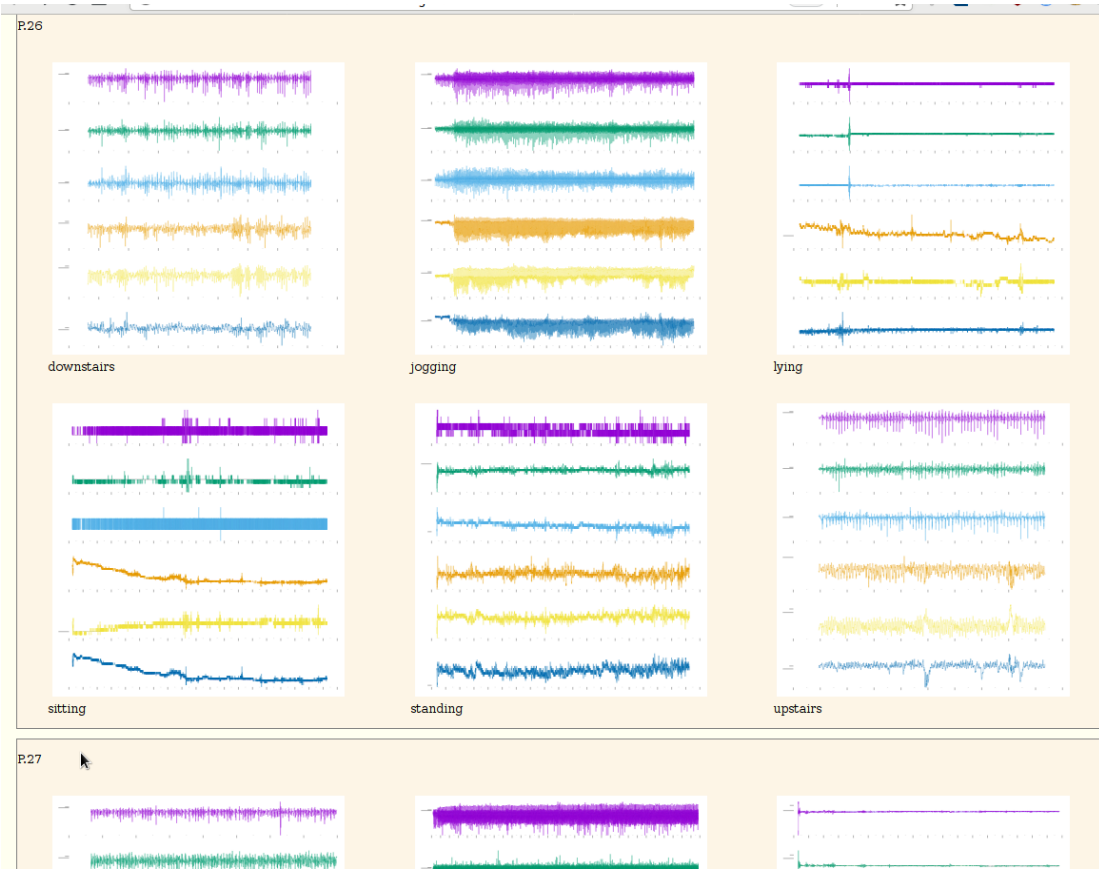


Fig. 10. A screen-shot of the ATV prototype displaying accelerometer traces.

been recent work on using signals from human voice to detect irregularities in properties such as pitch and timbre that may be used to predict the speakers emotional state [12].

Due to their low computational requirements, the methods of this paper may be used for on-board detection of physical activity using low-cost and energy efficient microprocessors in portable devices that are carried by a person. However, it is also frequently beneficial to upload the sensor data, in either raw or processed form, to data centers [13] where it may be recorded in a more durable form and subjected to more sophisticated and longer term analysis than is possible with portable devices. In the latter situation, especially when near real-time monitoring is required, wireless transmission of data may be achieved using a variety of networking technologies, such as WiFi and ZigBee [14]–[16].

This paper has focused on sensing of physical activities using sensors of mechanical properties (acceleration). In the broader context of using on-person sensors to monitor health status and activities, there is a very diverse and growing collection of sensing technologies for biomedical applications [17]. Another well studied category of sensors is optical. In particular, signals from cameras have been successfully used to gait detection and identification using biometric features extracted from the raw video [18]. In an even broader context, classification methods similar to those used by this work have been used in diverse applications including performance modeling of linear-system solvers [19]. Also related to sensing physical activities is the intuitively reverse task of using actuators to prompt physical motion. For instance, recent work has explored the use of functional electrical stimulation and vibrotactile feedback to teach motor skills [20].

VII. Conclusion

This paper has outlined the benefits of using data from body-attached accelerometers to determine the physical activities of individuals. It has provided a set of easily computed features that predict physical activity more accurately than the methods in prior work despite being simpler to compute. This paper has also briefly described the implementation of a prototype system that facilitates exploration of datasets of accelerometer traces in order to gain a better understanding of their characteristics.

Ongoing work is expanding the scope of the experimental study to investigate both whether higher accuracies may be achieved by augmenting the set of computed features and also whether features may be omitted without significantly degrading prediction accuracy. Another topic for ongoing and future work is evaluating these and similar methods on a diverse collection of accelerometer traces, from different participants, sensors (models, configurations), and attachments.

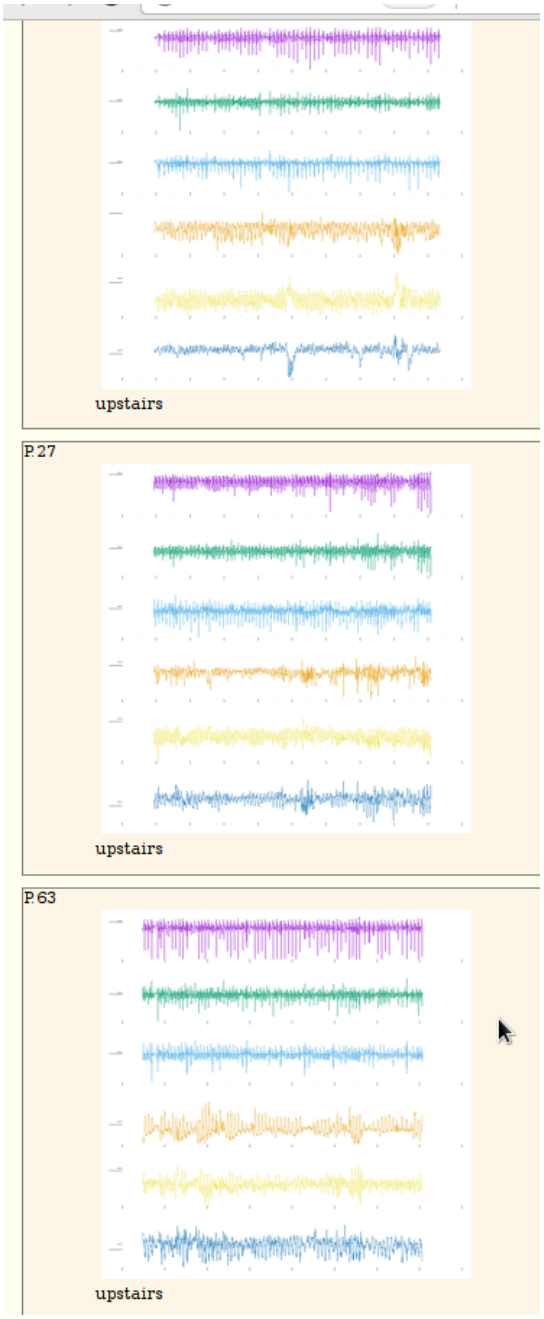


Fig. 11. A screen-shot of the ATV prototype being used to compare traces for the *upstairs* activity across participants.

ACKNOWLEDGMENTS

This work was supported in part by the US National Science Foundation grants 1848747 and others. It was enabled by the data provided by prior work [6] and benefited from the reviewers' comments.

REFERENCES

[1] Ewerton Naves Dias, Jos Vitor da Silva, Jos Lus Pais-Ribeiro, and Teresa Martins, "Validation of the advanced activities of daily living scale," *Geriatric Nursing*, 2018. http://www.sciencedirect.com/science/article/pii/S0197457218301733

[2] Cynthia M. Boyd, C. Seth Landefeld, Steven R. Counsell, Robert M. Palmer, Richard H. Fortinsky, Denise Kresevic, Christopher Burant, and Kenneth E. Covinsky, "Recovery of activities of daily living in older adults after hospitalization for acute medical illness," *Journal of the American Geriatrics Society*, vol. 56, no. 12, pp. 2171–2179, 2008. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1532-5415.2008.02023.x

[3] M. Powell Lawton and Elaine M. Brody, "Assessment of older people: Self-maintaining and instrumental activities of daily living," *Gerontologist*, vol. 9, pp. 179–186, 1969.

[4] Sudarshan S. Chawathe, "Recognizing activities of daily living using binary sensors," in *Proceedings of the IEEE International Conference on Universal Village (IEEE UV 2018)*. Cambridge, MA: MIT, Oct. 2018.

[5] ——, "Recognizing human falls and routine activities using accelerometers," in *Proceedings of the 9th IEEE Annual Computing and Communication Workshop and Conference (IEEE CCWC 2019)*, Las Vegas, Nevada, Jan. 2019.

[6] Sadiq Sani, Nirmalie Wiratunga, Stewart Massie, and Kay Cooper, "SELFBACK—activity recognition for self-management of low back pain," in *Research and Development in Intelligent Systems XXXIII*, Max Bramer and Miltos Petridis, Eds. Cham: Springer International Publishing, 2016, pp. 281–294.

[7] William W. Cohen, "Fast effective rule induction," in *Proceedings of the Twelfth International Conference on Machine Learning*. Morgan Kaufmann, 1995, pp. 115–123.

[8] Per Bothner, "Kawa: Compiling dynamic languages to the java VM," in *Proceedings of the USENIX Annual Technical Conference, FREENIX Track*. Berkeley, CA, USA: USENIX Association, Jun.15-19 1998, pp. 41–41. http://dl.acm.org/citation.cfm?id=1268256.1268297

[9] Ross Quinlan, *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann Publishers, 1993.

[10] Eibe Frank, Mark A. Hall, and Ian H. Witten, "The WEKA workbench." 2016, online appendix for Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, Fourth Edition, 2016.

[11] D. R. Hipp, D. Kennedy, and J. Mistachkin, "SQLite documentation," https://sqlite.org/docs.html, Jan. 2017, version 3.16.2.

[12] Natalie Ang, Doina Bein, David Dao, Luigi Sanchez, Justine Tran, and Navie Vurdien, "Emotional prosody analysis on human voices," in *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, Jan. 2018, pp. 737–741.

[13] Moises Levy and Jason O. Hallstrom, "A new approach to data center infrastructure monitoring and management (DCIMM)," in *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*, Jan. 2017, pp. 1–6.

[14] Reza Filsoof, Alison Bodine, Bob Gill, Stephen Makonin, and Robert Nicholson, "Transmitting patient vitals over a reliable zigbee mesh network," in *2014 IEEE Canada International Humanitarian Technology Conference - (IHTC)*, Jun. 2014, pp. 1–5.

[15] Srijita Chakraborty, Kaushik Das, Kaustavi Sen, Uttiyo Hari, Pritam Sarkar, Abhishek Raj, Malay Gangopadhyaya, Bob Gill, and Mrinmoy Chakraborty, "Elimination of higher order modes in wearable microstrip antenna," in *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Oct. 2016, pp. 1–4.

[16] Doina Bein and Bharat B. Madan, "Reducing the data communication delay in wireless sensor networks," in *2016 IEEE 12th International Conference on Intelligent Computer Communication and Processing (ICCP)*, Sep. 2016, pp. 361–368.

[17] Archita Bahattacharya, Malay Gangopadhyay, and Sankhyabrata Bandyopadhyay, "Designing of highly sensitive fiber bragg grating based volume refractive index sensors for biochemical applications," in *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Nov. 2018, pp. 918–923.

[18] Pratik Chattopadhyay, Shamik Sural, and Jayanta Mukherjee, "Information fusion from multiple cameras for gait-based re-identification and recognition," *IET Image Processing*, vol. 9, no. 11, pp. 969–976, 2015.

[19] Kanika Sood, Boyana Norris, and Elizabeth Jessup, "Comparative performance modeling of parallel preconditioned krylov methods," in *2017 IEEE 19th International Conference on High Performance Computing and Communications; IEEE 15th International Conference on Smart City; IEEE 3rd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, Dec. 2017, pp. 26–33.

[20] Abhishek Jha and Mrinal Sen, "Use of FES with vibrotactile feedback for motor learning," in *2016 International Conference on Internet of Things and Applications (IOTA)*, Jan. 2016, pp. 148–151.

A Novel Method for Mining Abnormal Expenses in Social Medical Insurance

Shengyao Zhou
School of Economics and Management
Beijing Jiaotong University
Beijing, China
zhoushengyao@bjtu.edu.cn

Runtong Zhang*, Senior Member, IEEE
School of Economics and Management
Beijing Jiaotong University
Beijing, China
rtzhang@bjtu.edu.cn

***Abstract*—With the continuous improvement of the medical insurance system and the continuous expansion of medical insurance coverage, the normal operation of the medical insurance fund has been closely related to the vital interests of the people. Medical fraud detection methods can effectively reduce the loss of medical insurance funds and improve medical quality. This paper proposes an effective fraud identification method to detect abnormal expenses in the medical insurance reimbursement process. Experiments on real medical insurance data prove that the method has good fraud detection accuracy. At the same time, as the amount of data increases, the performance advantages of our proposed method are more obvious.**

Keywords—medical insurance, abnormal detection, local outlier factor

I. INTRODUCTION

With the continuous improvement of people's living standards, citizens' own social security awareness has been continuously improved. The medical insurance system occupies an important position in the social security system with its indispensable importance. Since the implementation of the system, the medical insurance fund has played an important role in solving the problem of people seeking medical treatment.

Medical insurance is considered to be a tripartite process of game between the insured, medical institutions and medical insurance providers. The driving of the interests of the three parties is often easy to induce fraud. From the perspective of hospitals, the purpose of increasing hospital income is often achieved through over-diagnosis and treatment, over-use of drugs, and decomposition of hospitalization; From the perspective of insured, fake medical treatment, frequent consultations and illegal reimbursement have become a means to reduce personal medical expenses; In terms of medical insurance providers, there are occasions where the review of medical insurance settlement data is insufficient. Under these circumstances, the intensified medical fraud has caused huge losses to the medical insurance fund and threatened the normal operation of the medical insurance fund. Therefore, in order to achieve the long-term normal operation of the medical insurance fund, it is essential to increase the detection of medical fraud.

At present, most of the review of medical insurance violations are mainly solved manually, which is extremely heavy, inefficient, and does not allow for good detection of healthcare fraud. With the widespread application of data mining and machine learning in the medical field, the need to use data mining and other technologies to solve medical fraud has emerged. Some scholars began to apply data mining technology to the detection of medical insurance fraud. For example, Shi [1] used the semi-supervised Isomap method and the Local Outlier Factor (LOF) detection method to identify the cost of abnormalities, but did not take into account the individual differences of patients when selecting features. Xie [2] combined information entropy and LOF methods for abnormal fund expense detection, and also did not take into account the characteristics of the patients themselves. Since most of the current researches are aimed at a certain kind of fraud, this paper combines the clustering algorithm with the LOF method, selects the specified common diseases, and proposes a new medical expense abnormality detection (MEAD) method and an improved medical expense abnormality detection (IMEAD) method to help reduce the workload of audits.

The rest of the paper is organized as follows: Section 2 reviews related works on fraud detection issues; Section 3 gives the relevant definitions and introduces the MEAD method and IMEAD method in detail; Section 4 uses real medical insurance data for experimental analysis and discussion; Section 5 summarizes our work and discusses several future research directions.

II. RELATED WORK

Medical insurance fraud is not a problem unique to a country, and countries around the world that implement medical insurance systems are facing corresponding problems. Medical insurance fraud will inevitably cause unpredictable losses to the medical insurance fund and there is widespread concern about how to reduce the incidence of fraud. At present, the research on medical insurance fraud is mainly divided into three aspects: the causes and characteristics of fraud, how to combat fraud and the identification of fraud.

This work was supported in part by the National Social Science Foundation of China (Major Project) under Grant 18ZDA086, in part by the Fundamental Research Funds for the Central Universities under Grant B19JB500230, and in part by the Beijing Logistics Informatics Research Base.
*Corresponding author: Prof. Runtong Zhang, email: rtzhang@bjtu.edu.cn.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

In terms of research on the causes of fraud and its characteristics, Skiba [3] systematically analyzed the causes of fraud by summarizing the experience of practitioners related to anti-fraud in medical insurance. Dionne [4] analyzed the possible risks of fraud based on asymmetric information theory and explained the causes of fraud from the fraudster's perspective. Krause [5] focused on the patient and analyzed the potential characteristics of various fraudulent behaviors. Kathryn [6] took Australian private companies as the research object and analyzed the impact of medical insurance-related policies on the occurrence of fraud and fund operation. At the same time, it points out that the occurrence of insurance fraud has common characteristics.

In terms of anti-fraud behavior research, Othman [7] surveyed population samples in the form of structured questionnaires, and pointed out that setting up fraud hotlines and legal designers are the most effective fraud detection and prevention mechanism in the public sector. Rudman and William [8] analyzed the fraud in the US medical insurance, and put forward medical insurance anti-fraud suggestions from the perspectives of strengthening medical insurance information management, improving medical insurance related systems, and strengthening the education of insured persons and medical-related practitioners. Sun [9] summarized the anti-fraud mechanism of medical insurance in European and American countries, and pointed out that it is necessary to strengthen the restriction and supervision of medical insurance institutions, and make full use of modern information technology to improve the level of anti-fraud.

In terms of fraud detection research, traditional detection methods are mainly rule-based methods [10-11]. Once medical treatment violates the rules established by experts, it is judged as fraud. The effectiveness of this method is restricted by the correctness of the rules. With the widespread application of big data technology in the medical field, data mining algorithms are gradually being applied to the detection of medical fraud. The data mining algorithm regards the fraud in the medical treatment process as abnormal data records, and converts the fraud detection problem into the classic outlier detection problem. Common outlier detection methods can be divided into four main categories: statistics-based, classification, distance, and clustering.

The classification-based anomaly detection method [12] divides the data into two types, normal and abnormal, and uses the labeled data for training to convert the anomaly detection problem into a two-classification problem. The basic idea of the distance-based anomaly detection method is based on the distance between data points. The abnormal point will be greatly different from its neighboring points, and the normal point will be very close to its neighboring points. Different methods use different distance or similarity measurement methods. There are two main ways to measure the distance between anomalous points and its neighboring points. One is to calculate the distance between each data point and the set of neighboring points, and then judge the anomaly by distance ; The other is to calculate the density of each region in the data, and treat the data in the low-density region as abnormal. Breunig [13] defines a local outlier factor for each data point. The larger the value of the local outlier factor, the more likely it is to be an outlier. Statistics-based anomaly detection methods [14-15] assume that normal data are generated by a certain statistical model, and those that do not conform to the model's distribution law are regarded as abnormal data. This method highly depends on the model assumptions of the data. The cluster-based anomaly detection method [16-18] believes that normal points must belong to clusters containing more data points, while abnormal points belong to clusters containing fewer data points or do not belong to any cluster. This method first clusters, and then determines whether an abnormal data point is based on the size of the cluster.

In summary, data mining algorithms are beginning to be applied to the detection of medical fraud, which greatly reduces the overhead of manual auditing, but most of the algorithms have poor interpretability. Since statistical methods are highly dependent on the assumption of data distribution, and classification-based detection methods require a large amount of labeled data, these two methods are generally not suitable for detection scenarios of medical insurance fraud. Therefore, on the basis of previous research, this paper proposes the MEAD method and the IMEAD method based on the LOF algorithm, and applies them to actual data analysis.

III. Materials and Methods

The MEAD method is a combination of the clustering algorithm and the LOF method idea, which first uses the clustering algorithm to gather the datasets into clusters, and then divides the large and small clusters according to the cluster size and calculates the outlier factor of the data in each cluster respectively. Usually, the outlier factor of data points in a large cluster will be low, while the outlier factor of data points in a small cluster will be relatively high. In the same cluster, the data points at the edge of the cluster are more likely to be abnormal points, and the points at the center of the cluster are more likely to be normal points. In this paper, we use MEAD algorithm to detect abnormal medical expenses, calculate the *lof* value for each data, and finally select a small part of the data as the candidate abnormal record set according to the *lof* value.

A. Problem Definition

In order to facilitate the subsequent description of the specific steps of the fraud detection method, this section introduces the symbols used in the paperand their related definitions.

Definition 1: Let d be the medical insurance data record, the medical insurance dataset MID contains m data records, then $MID = \{d_1, d_2, \dots d_m\}$. Let A be the attribute of d, and each piece of data contains n attributes, then $d = \{A_1, A_2, \dots A_n\}$.

Definition 2: Let C be the clusters that the medical insurance dataset is divided into after clustering, and k is the number of clusters, then $MID = \{C_1, C_2, \dots C_k\}$. Let λ be the threshold 1, which is used to distinguish the size of the clusters, where large cluster is marked as LC, small cluster is marked as SC, then

$$C_i = \begin{cases} C_i \in LC, |C_i| > \lambda m \\ C_i \in SC, |C_i| \leq \lambda m \end{cases} (1 \leq i \leq k)$$

Definition 3: If a data record belongs to a cluster marked LC after clustering, its *lof* is determined by the size of the cluster and the distance from it to the center of the cluster; if a data record belongs to the cluster marked SC after clustering, the *lof*

is determined by the size of the cluster and the distance between it and the center of the nearest large cluster. The size of *lof* indicates the degree of abnormality of the data record. The larger the *lof* value, the greater the possibility that the data is an abnormal record.

$$lofd_s = \begin{cases} \frac{1}{|C_i|} \cdot min\left(dist(d_s, C_j)\right), d_s \in C_i, C_i \in SC, C_j \in LC \\ \frac{1}{|C_i|} \cdot \left(dist(d_s, C_i)\right), \quad d_s \in C_i, C_i \in LC \end{cases}$$
$$(1 \le i \le k, 1 \le j \le k, 1 \le s \le m)$$

Definition 4: Let *N* be a counter, and is μ be a threshold 2. When a new data record is added, if it belongs to the cluster marked *SC* after clustering, the *lof* value of each record is update immediately; If it belongs to the cluster marked as *LC*, the *N* count is increased by 1, and when the ratio of *N* to the current cluster size exceeds the threshold μ, the *lof* value of each data is update immediately.

$$lofd_s = \begin{cases} lofd_s \cdot \frac{|C_i|+1}{|C_i|}, & C_i \in SC \\ lofd_s \cdot \frac{|C_i|+N}{|C_i|}, & C_i \in LC, \frac{N}{|C_i|} > \mu \end{cases}$$
$$(1 \le i \le k, 1 \le s \le m)$$

B. Medical Expenses Abnormal Detection Method

The MEAD method first performs cluster analysis on the dataset, divides large clusters and small clusters according to the clustering results, and then calculates the *lof* value for each record. Considering that the Squeezer algorithm does not need to specify the number of clusters initially, it can give a more natural clustering effect. Therefore, the Squeezer method is selected for clustering in this article. For the first record, create a new cluster through *addNewCluster*. For each subsequent record, use *calculateSimilarity* to judge its similarity with the current existing cluster, and find the most similar cluster and the largest similarity. If the maximum similarity exceeds the preset threshold, use *addToCluster* to add the record to the corresponding cluster, otherwise, use *addNewCluster* to create a new cluster.

After the clustering work is completed, the *lof* value of each data record needs to be calculated. For a given threshold λ, it is the basis for dividing large and small clusters. For the results of the first step clustering, the large clusters and small clusters are divided according to definition 2. The large clusters are marked as *LC* and the small clusters are marked as *SC*. Traverse the dataset again, for each data record, according to definition 3, if the record belongs to the cluster marked as *LC*, use *calculateLClof* to calculate the *lof* value of the record; if it belongs to the cluster marked as *SC*, use *calculateSClof* to calculate the record *lof* value, the specific algorithm flow is shown in Algorithm 1

C. Improved Medical Expenses Abnormal Detection Method

As new data are coming into the dataset successively, if the MEAD method is used for fraud cost detection, the model needs to be retrained each time, which is a big overhead when the

Algorithm 1. MEAD

Input: *MID*, λ, *C*, *maxsim*, *m*
Output: *SC*, *LC*, $lofd_s$
1) **for** d_s in *MID* **do**
2) **if** *s*==1 **then**
3) *addNewCluster*(d_s)
4) **else**
5) **for** C_i in *C* **do**
6) *calculateSimilarity*(d_s, C_i)
7) **return**(*cursimilarity*, *index*)
8) **end for**
9) **end if**
10) **if** *cursimilarity* > *maxsim* **then**
11) *addToCluster*(d_s, *index*)
12) **else**
13) addNewCluster(d_s)
14) **end if**
15) **end for**
16) **for** C_i in *C* **do**
17) **if** $|C_i| > \lambda m$ **then**
18) $LC \coloneqq LC \cup C_i$
19) **else**
20) $SC \coloneqq SC \cup C_i$
21) **end if**
22) **end for**
23) **for** d_s in *MID* **do**
24) **if** $d_s \in C_i$ and $C_i \in SC$ **then**
25) $lofd_s = calculateSClof(d_s, C)$
26) **else**
27) $lofd_s = calculateLClof(d_s, C_i)$
28) **end if**
29) **end for**

Algorithm 2. IMEAD

Input: *MID'*, μ, *C*
Output: $lofd'_s$
1) **for** d'_s in *MID'* **do**
2) *Squeezer*(d'_s)
3) **return**(C_i, *index*)
4) **if** $C_i \in SC$ **then**
5) *addToCluster*(d'_s, *index*)
6) **for** d'_s in C_i **do**
7) $lofd'_s = cupdateSClof(d'_s, index)$
8) **end for**
9) **else**
10) $N \coloneqq N + 1$
11) $lofd'_s = calculateLClof(d'_s, index)$
12) **if** $N > \mu|C_i|$ **then**
13) **for** d'_s in C_i **do**
14) $lofd'_s = updateLClof(d'_s, index)$
15) **end for**
16) $N = 0$
17) **end if**
18) **end if**
19) **end for**

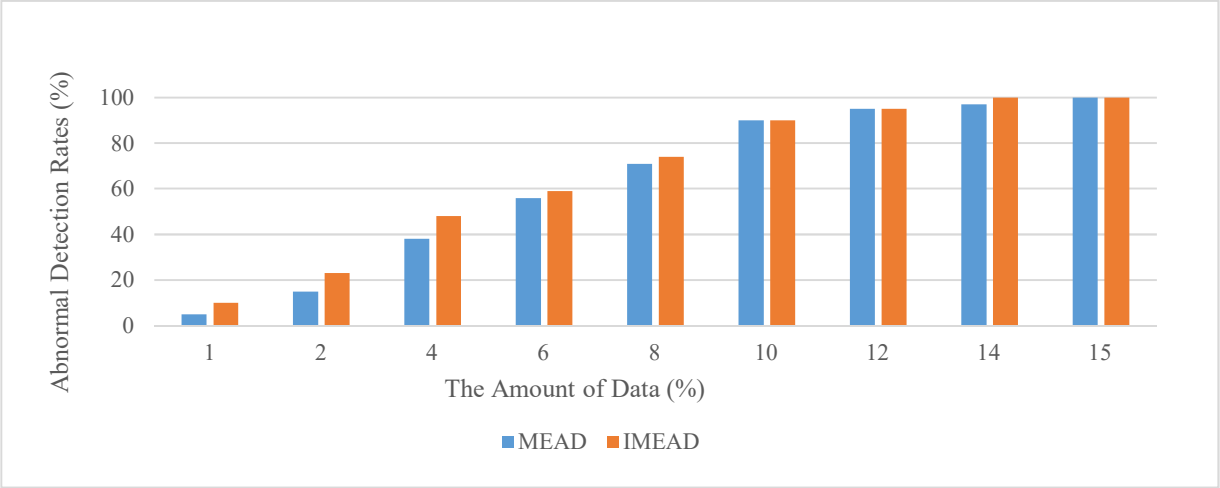


Fig. 1. Abnormal detection rate comparison

TABLE I. PART OF THE DATA ITEM INFORMATION

Type of Data	*Data Item*
Personal Information	Medical Card Number
	Sex
	Age
	Disease Code
Hospitalization Information	Hospitalization Number
	Medical Card Number
	Date of Admission
	Date of Discharge
	Types of Discharged Diseases
	Total Cost
	Medicine Fees
	Operation Fees
Payment Information	Hospitalization number
	Medical Card Number
	Disease Code
	Total Cost
	Actual Compensation Amount
	Out-of-pocket Amount

amount of data is large. Therefore, we can use the IMEAD algorithm, which is based on the idea that the model is first trained using the MEAD method for existing data until the model is relatively stable. When new data arrives, use the Squeezer algorithm to cluster and judge each record. If the record belongs to the cluster marked as *SC*, use *addToCluster* to add the record to the current cluster, and according to definition 4, use *updateSClof* update the *lof* value of all data points in the current cluster; if the record belongs to the cluster marked as *LC*, set the counter *N*, and add 1 to the count *N*, use *calculateLClof* to calculate the *lof* value of the data. Only when the ratio of N to the current cluster size exceeds the threshold μ, *updateLClof* is used to update the *lof* values of all data points in the current cluster, which can greatly reduce the overhead of frequently updating *lof* values. The specific algorithm flow is shown in Algorithm 2.

IV. EXPERIMENT RESULTS

In this section, we use the proposed method for experimental analysis. First, we describe the experimental environment and

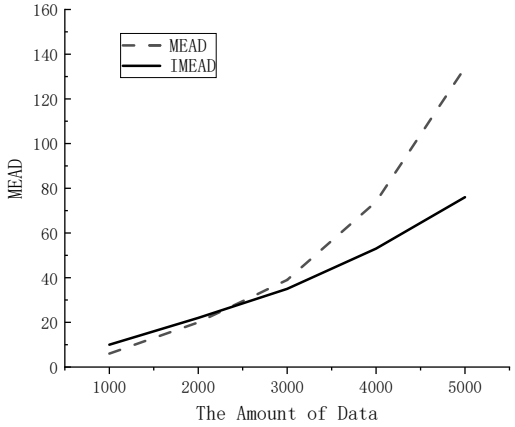


Fig. 2. Time efficiency comparison

provide implementation details. Then, we proved the effectiveness of the method by comparing the abnormal detection rate and time efficiency.

A. Experimental Settings

Our experimental platform is Windows Sever 2008 R2 Datacenter, CPU 3.30 GHz, 8 GB of memory, implemented in Java on Eclipse 4.5. The experiment used a real dataset of medical reimbursement in a county from 2013 to 2016, including 391,185 medical insurance records. Table 1 shows part of the data item information of the dataset.

Through the statistical summary of the dataset by disease type, it is found that the total cost of cardiovascular and cerebrovascular diseases is the highest. Therefore, this article focuses on fraud detection for the hospitalization records of patients with cardiovascular and cerebrovascular diseases with a total of 71046 data, and the top 5%-10% of the final *lof* values were selected as candidate outliers.

B. Results and Analysis

For the hospitalization data of patients with cardiovascular and cerebrovascular diseases, we used the MEAD and IMEAD methods to detect the *lof* value of each data, respectively, and then sorted them in descending order according to the *lof* value, and selected the higher *lof* value as the candidate for the outlier. Figure 1 shows a comparison of abnormal detection rate using MEAD algorithm and IMEAD algorithm. It can be seen from Figure 1 that the anomaly detection rate of MEAD and IMEAD is increasing as the amount of data increases. The anomaly detection rates of both have exceeded 50% for the top 6% of the data volume. In addition, IMEAD has a higher abnormal detection rate than MEAD when the amount of data is small, and finally, under the top 14% of the data volume, its abnormal detection rate reached 100%; In contrast, MEAD finally reached 100% of the abnormal detection rate under the top 15% of the data volume.

The time efficiency of the two algorithms is shown in Figure 2. It can be seen from Figure 2 that the time efficiency of the MEAD method is slightly higher the IMEAD method in the case of fewer data records, but as the amount of data increases, the performance advantage of the IMEAD method becomes more obvious. This is because the IMEAD method uses the original dataset to train the model until it reaches a stable condition. If new data is added, there is no need to retrain the model, but to directly determine the cluster to which the new data belongs, reducing the time overhead of retraining the model. At the same time, only when a record belongs to a cluster marked as *SC*, the *lof* value of all records in the cluster is updated immediately; if it belongs to a cluster marked as *LC*, the *lof* value is updated only when the specified threshold value is reached, which greatly reduces the time overhead of updating the *lof* value.

V. Conclusion

The normal operation of the medical insurance fund is very important. In this paper, we proposed an effective fraud identification method. The MEAD method combines the clustering algorithm with the LOF method to detect abnormal expenses in the medical insurance reimbursement process. In addition, for the newly added data, we proposed the IMEAD method to improve the efficiency of the algorithm. We use a real dataset from the medical insurance system of a county in China, which contains 390,000 pieces of medical insurance data. Experimental results show that the MEAD method and IMEAD method have good fraud detection accuracy. At the same time, as the amount of data increases, the performance advantage of the IMEAD method becomes more obvious.

At present, this method has been integrated into the medical data analysis platform and has provided decision support for the auditors in the medical insurance claims system to assess the possibility of fraud. In subsequent research, we will focus on the differences between different diseases and explore the potential connection between disease types and medical behaviors in conjunction with the medical insurance knowledge base.

References

[1] SHI Y, SUN C, LI Q, et al, "A fraud resilient medical insurance claim system," in the thirtieth AAAI Conference on Artificial Intelligence. USA: AAAI Press, 2016, pp. 4393–4394.

[2] Xie Z. P., Li X. Y., Wu W. Y., et al, "An improved outlier detection algorithm to medical insurance," in IDEAL, 2016, pp. 436–333.

[3] Skiba J. M., Disch W. B., "A phenomenological study of the barriers and challenges facing insurance fraud investigators," Journal of Insurance Regulation, vol. 33, no.4, 2014, pp. 87–114.

[4] Dionne, Georges, Gagné, et al, "Replacement cost endorsement and opportunistic fraud in automobile insurance," Journal of Risk and Uncertainty, vol. 24, no.3, 2002, pp. 213–230.

[5] Krause H. J., "A patient-centered approach to health care fraud recovery," The Journal of Criminal Law and Criminology, vol. 96, no. 2, 2006, pp. 579–619.

[6] Flynn K., "Financial fraud in the private health insurance sector in Australia," Journal of Financial Crime, vol. 23, no. 1, 2015, pp. 143–158.

[7] Othman R., Aris N. A., Mardziyah A., et al, "Fraud detection and prevention methods in the malaysian public sector: accountants' and internal auditors' perceptions," in International Conference on Financial Criminology 2015, 2015, pp. 59–67.

[8] Rudman W. J., Eberhardt J. S., Pierce W., et al, "Healthcare fraud and abuse," Perspectives in Health Information Management, vol. 6, no. 2, 2009, pp. 111–112.

[9] J. Sun, Y. Gan, "Medical insurance anti-fraud mechanism from the perspective of cooperative governance: international experience and enlightenment," China Health Policy Research, vol. 10, no. 10, 2017, pp. 28–34.

[10] Milo T., Novgorodov S., Tan W., "Rudolf: interactive rule refinement systemfor fraud detection," PVLDB, vol. 9, no. 13, 2016, pp.1465 –1468.

[11] Subelj L., Furlan S., Bajec M., "An expert system for detecting automobile insurance fraud using social network analysis," Expert Systems with Applications, vol. 38, no. 1, 2011, pp.1039–1052.

[12] Roberts S. J., Penny W., Pillot D., "Novelty, confidence and errors in connectionist systems," in Intelligent Sensors. [S.l.]: IET, 1996, pp. 10/1-10/6

[13] Breunig M. M., Kriegel H. P., Ng R. T., et al, "Optics-of: identifying local outliers," Lecture Notes in Computer Science, 1999, pp. 262-270.

[14] H. Huang, J. Lin, C. Chen, "Overview of outlier data mining," Computer Application Research, vol. 23, no. 8, 2006, pp. 8–13.

[15] Liu B., Yin J., Xiao Y., et al, "Exploiting local data uncertainty to boost global outlier detection," in IEEE International Conference on Data Mining. [S.l.]: IEEE Computer Society, 2010, pp. 304–313.

[16] Ester M., Kriegel H. P., Xu X., "A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise," in International Conference on Knowledge Discovery and Data Mining. USA: AAAI Press, 1996: pp. 226–231.

[17] Ng R. T., Han J., "Efficient and effective clustering methods for spatial data mining," in International Conference on Very Large Data Bases. San Francisco: Morgan Kaufmann, 1994, pp. 144–155.

[18] Zhang T., Ramakrishnan R., Livny M., "Birch: An efficient data clustering method for very large databases," ACM SIGMOD Record, vol. 25, no. 2, 1996, pp. 103–114.

[19] He Z., Xu X., Deng S., "Squeezer: an efficient algorithm for clustering categorical data," Journal of Computer Science and Technology, vol. 17, no. 5, 2002, pp. 611–624.

Technological and Medical Progress in Rural Bangladesh; A Survey and Study of the Challenges

Navid Bin Ahmed
Independent University Bangladesh (IUB)
Dhaka, Bangladesh
navid_bin_ahmed@yahoo.com

Shahriar Khan
Independent University Bangladesh (IUB)
Dhaka, Bangladesh
skhan@iub.edu.bd

***Abstract*— While there has been great progress in the technology of medical devices, most people in developing countries do not have access to the healthcare technologies. Through a survey of households in rural Bangladesh, some challenges were identified in accessibility of technology and medical technology for the residents. There was just one doctor in the vicinity, but no hospital or clinic. Most took medicine based on the advice of their local store pharmacists. Nearly all were found to be living day to day, on only one main source of income. Women's empowerment had progressed, but education of women was very low.**

Keywords—developing, Bangaldesh, electronics, challenges, survey, medical, technological, socio-economic, scientific, advancement, global etc.

I. INTRODUCTION

While much research has been undertaken on the advancement of healthcare and technologies in developing countries, less work has been conducted on the limitations of their availability, accessibility and affordability [1]-[3]. This paper seeks reasons behind the lack of accessibility of technology and medical technology in the country.

In the past, there have been attempts to manufacture affordable technologies for the masses. The authors have developed low-cost healthcare electronics for developing countries [4]-[7]. This approach of the authors is expected to promote both smart medical devices, and profitability to investors. The rural people of Bangladesh need to be slightly more affluent before they can access smart devices and medical equipment of today, various IP technologies, and nationwide telemedicine services in collaboration with telecommunication operator Grameenphone in Bangladesh [8]-[15]. Developments in the power sector have largely allowed the electrification of most villages [16]- [19].

To ensure affordability, even cheaper technologies, such as for medical and wellness devices may be researched [20]-[21]. Improving the literacy rate [22], and facilitating women's empowerment [23] are expected to improve the economy and bring about greater women's empowerment [23].

A systematic survey was conducted on technological and medical progress in rural Bangladesh. The paper seeks reasons behind the lack of accessibility of technology and medical services in the country.

Several challenges to do with quality of life were found. Correlation was found between residents' occupation, financial status, and lifestyles

II. METHODOLOGY

Salandar is a village in Thakurgaon district of Bangladesh where 0.60 million people live in an area of 264 square miles.

Five houses were chosen for the survey, marked by unique numbers, 82306, 82428, 82425, 82422 and 82445. The survey covered areas (*Para* or *Moholla*) of the village named *Danispara*, and *Panjiarpara*. In these 5 households, 11 people were adults and 10 were below eighteen years of age.

Questions asked were on their finances, use of technology, health status and available medical facilities. Other questions included their name, age, occupation, monthly income, medical conditions, financial status (family income, expense, and savings), and the property they owned.

MS Excel was used to illustrate the graphs and figures in this paper.

III. USE OF TECHNOLOGY

All houses were found to have electricity. Their use of technology ranged from waking up to the alarm clock in the morning to chatting on *Facebook* late at night. Other than the smart phone, electronic appliances were almost unknown to them. About 25% could afford a few essential electronic products (figure 1).

None had a computer, radio or a refrigerator. Only 4% could afford a smart phone.

IV. HEALTH AND ACCESS TO HEALTHCARE

The second part of the survey was about their health and access to healthcare. The finding was that they had poor access.

Six out of the eleven adults or nearly 55% of the total said that their health condition was good, while almost 45% believed that theirs was moderate.. No one believed that they were physically unfit or weak.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

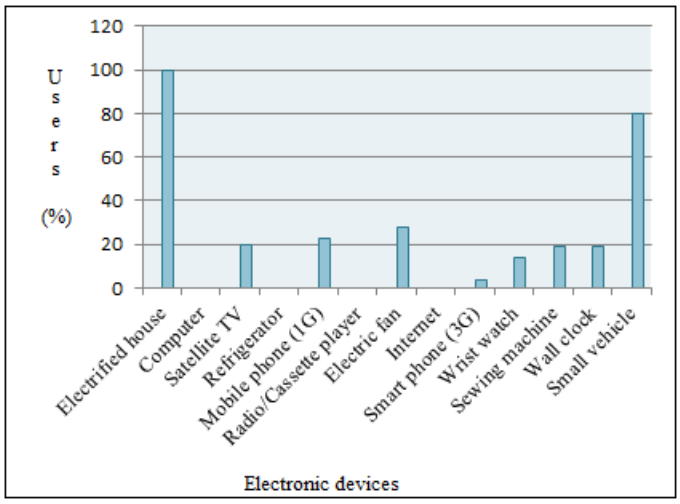


Fig. 1. Users of electronic accessories

Common diseases and disorders included fever, headache, colds and coughs (figure 2). Almost none was found to have any serious illness. One of the two members in a household (HHID: 82425) was injured the previous year and the other one became ill due to a poisonous insect bite. Both took medicine to recover.

The medical facilities in most of the village were poor. There was only one qualified doctor in the whole area, who worked far away in the District/ *Sadar* hospital. People mostly depended on the local pharmacists for everyday treatment. Those in need of hospitalization got admitted to the District/ *Sadar* hospital.

Most believed they had to remain fit to work to maintain the family. Every family was found to have purchased some medicine at one time or another. One family was even found

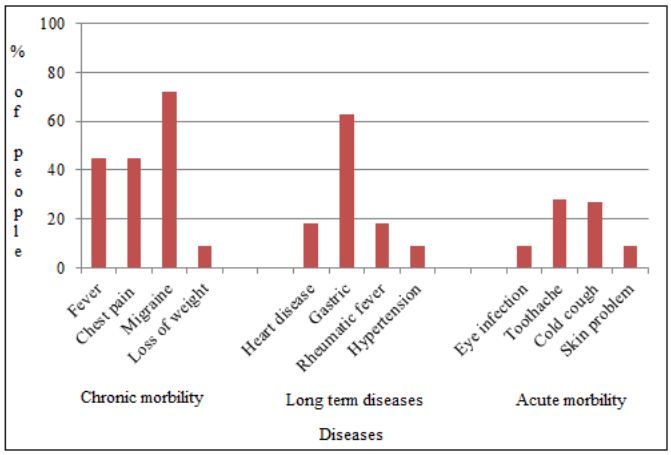


Fig.2. Common diseases and disorders

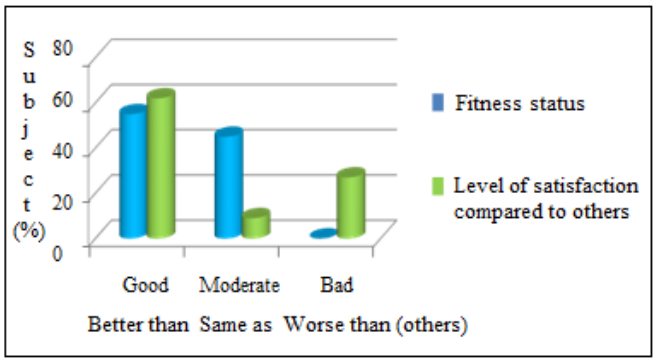


Fig. 3. Peoples' satisfaction level over health and medical condition

to be relying on their neighbors for medical expenses. Table I shows the medical expenses of the previous month.

The drugs commonly taken were *Napa Extra, Neotek, Omiprazol, Methox, Endomet, Avicof, Offcof, Omep, Vargon, Exiam, Omniax, Excer, Pantobex, Amylin, Abecab, Zelazin* etc. From their responses to questions on psychology and mental health, some showed signs of depression. Depression appeared more common among those without a source of income.

Most believed that their health and medical facilities were adequate. About 60% of respondents found their present health condition better than average, whereas nearly 40% believed that theirs was the same as and worse than others (fig. 3).

TABLE I. AVERAGE MONTHLY MEDICAL EXPENSE

HHID	Expense (BDT)
82306	70
82428	160
82425	350
82422	70
82445	Free purchase

IV.

V. FAMILY FINANCES

A. Poor Family Financial Condition

3 out of the 5 families surveyed were involved in agricultural work. Their financial situation is given in Table II:

TABLE II. FAMILY FINANCES

HHID	Member	Adult	Earning Member	Monthly Income (BDT)	Expense (BDT)
82445	5	2	2	4,500	4,000
82422	5	2	1	7,500	6,500
82425	4	3	2	5,000	4,500
82428	4	2	2	15,000	12,000
82306	3	2	1	7,500	5,500

The main two occupations were farming and household work. Only around 10-15% of the households could save a small amount at the end of the month. Very few families had more than one source of income. Most were forced (unsuccessfully) to look for a second one.

In case of shortage of work, people tended to migrate to the cities, often for jobs at a garments factory. In a house of five, the household head, along with his wife, had moved for work to a garment factory in the capital city (Household ID: 82445).

Except for the land they lived on, almost no family owned any farming land. They lived on someone else's land for farming or other agricultural work.

Regardless of the hurdles, most managed to earn enough for survival. Their monthly gross income hardly crossed 10,000 BDT (approximately 117 US$). The monthly income and expenses are shown in figure 4.

B. *Lower Women's Literacy and Empowerment Rate*

Women were involved in various economic activities. In many households, wives and teenage daughters contributed to the family income. Some women tailored clothes in their spare time. Some women ran a small business, such as a tea stall, side by side with their husbands.

Out of 13 women, 7 were adults (above 18), 3 of whom worked for income. Mrs. *Majeda Begum* (33 yrs) made an income from cultivable land she inherited earlier from her parents. Ms. *Munni Akter* (23 years) earned money by tailoring. Ms *Zolekha* (30 years) migrated to Dhaka along with her husband to work in a garment factory. Her mother could not tell us the exact amount she earned. All 6 of the rest of the women

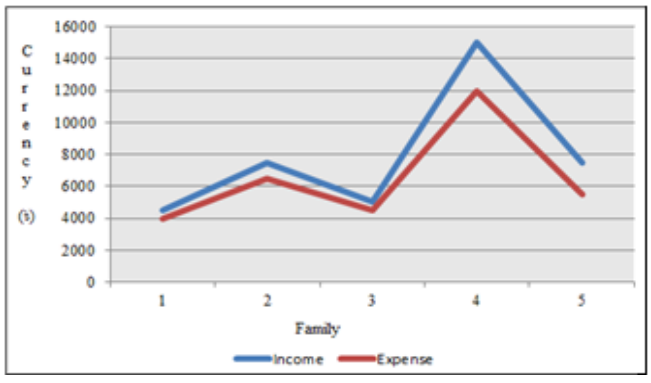


Fig. 4. The minimal ratio of monthly expense to income

were at schools and colleges, and were around 18 years of age.

More than one-fourth of the adult women or around 31% of them had no institutional education, while the majority of the women were primary and high school drop-outs (31% and 15% respectively). 23% of women received secondary education (figure 5). Hence, the rate of women being self-empowered is very poor at around 25%. This is even more concerning compared to the total (male and female people) which is only 14%.

The women were in need of an environment which is friendly, supportive and secure (Table III). Women above 18 years old, with primary and/or secondary education, have been focused on to identify their level of contribution. Keeping the female students out of the count, we found that around 30% of women had no economic contribution, while only 23% (3 women) added to their family incomes. Students were nearly 47% and almost 30% did not contribute financially. Based on their age, education and occupation, we further analyzed the economic contribution that each woman played.

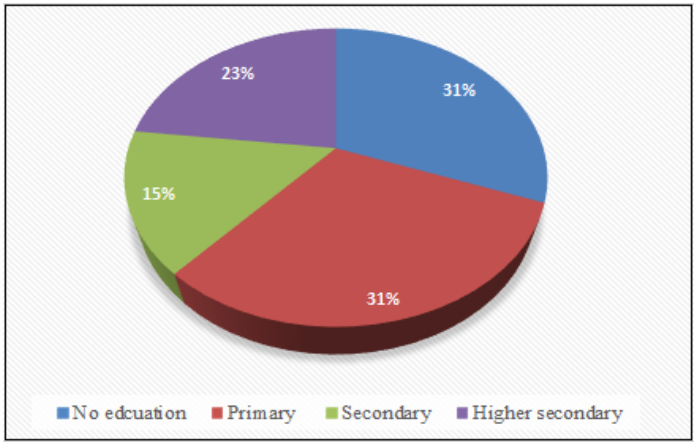


Fig. 5. Female literacy rate

TABLE III. ECONOMIC CONTRIBUTION OF WOMEN

Subject	Age (years)	Occupation	Income/Month (BDT)
1	33	Agribusiness	1,000
2	15	Student	N/A
3	51	Housewife	no contribution
4	23	Stitching	500
5	17	Student	N/A
6	30	Garment worker	2,500
7	55	Housewife	no contribution
8	4	Student	N/A
9	23	Housewife	no contribution
10	4	Student	N/A
11	33	Housewife	no contribution
12	11	Student	N/A
13	6	Student	N/A

C. *Limited Food Availability*

13 out of every 21 people or about 3 of every 5 families directly depended on agriculture, which is 60% of the total (households surveyed). The availability of a sufficient amount of food in appropriate quantity and quality is ensured from domestic production, sometimes from imports if needed.

The main occupations of the people of *Salandar* were cattle breeding, farming and household work. Almost no one could store food for months. Their existence was mostly hand-to-mouth. When they did not have land for farming, or if there was a shortage of work, they looked for other sources of activities. The survey found people from 2 out of 5 households starving for want of food (Table IV).

TABLE. IV REASON FOR FOOD SCARCITY.

HHID	HH Head and Wife	Reason for Food Scarcity
82306	Mr, *Profullo Chandra Roy* and Mrs. *Loliti Rani*	Previously the house head used to earn much. Due to illness, he can hardly work.
82445	Mr. *Habulu* and Mrs. *Zolekha*	A large family consisting of five members. House head shifted to Dhaka with his wife. His mother sometimes starves with two grandchildren.

Only one family (20% of total) was found to be comparatively well-off. They earned well and managed to save. The comparison of income against food cost per month is very poor, represented in Fig. 6.

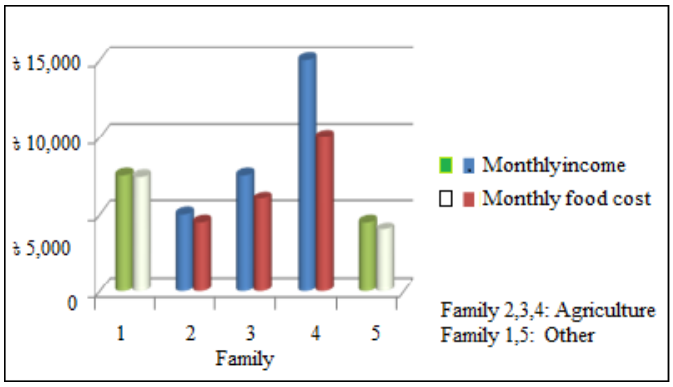


Fig. 6. Monthly family income and cost of food

VI. HOPES AND INSPIRATIONS

The psychology of the people including their thoughts, emotions and behavior were closely tied with their earnings.

A. *Simplicity*

The residents were found to be psychologically less complicated than city dwellers. When we knocked at their doors, their response was polite, even though they knew their lives would change little from these surveys.

B. *Selflessness*

Residents had few luxuries in life, yet appeared to sleep at night in peace. They found happiness in simple things, such as warming themselves at an open fire.

C. *Sense of Hospitality*

They were mostly less educated, but they welcomed and treated the surveyors hospitably. They treated us to a meal they had already prepared.

They were found to live in extended families in peace and harmony.

They could smile even when starving for days. In spite of a shortage of wealth and food, there was self-contentment.

VII. CONCLUSION

This is a study of technical and medical progress at *Salandar* in *Thakurgaon* of Bangladesh. On average, only about 25% of people could afford essential scientific and technological benefits. There was only one certified doctor with no adequate hospitals nearby. Agriculture was the leading occupation. Ensuring quality education and female empowerment may play key roles in bringing changes in individuals, families and society.

The development of a country is determined by income and standards of living. For proper implementation of a plan of development, the rural masses, such as the people of *Salandar* must be kept in consideration.

The World Bank, ADB, UNESCO and some NGOs contribute to the welfare of these people. But based on our work, it is clear that these efforts are too insignificant to cover rural masses. These people should be made aware of the importance of education, good health and standard of life.

REFERENCES

[1]. S. Xiaowei, Y. Hui, L. Mingxuan, C. Jiahui, Z. Xianzeng, "Current status and progress on virtual reality technology in medical field", Laser and optoelectronic progress, 57(1), January 2020.

[2]. L. Hakizimana, M. Wannous, "Distributed medical record system using portable devices, case study of Rwanda work in progress", IEEE International Conference on Applied System Invention (ICASI), IEEE Press, New York, 2018.

[3]. L. J. Frost, M. R. Reich, "Creating access to health technologies in poor countries", Health Affairs, 28(4), pp 963-973, July 2009.

[4]. N. B. Ahmed, S. Khan, K. N. Tahsin, and I. Afrida, "Implementation of a low-cost heartbeat monitor as part of a larger health monitoring system". 3rd International Conference on Inventive Computation Technologies, IEEE, pp. 813-818, 2018.

[5]. N. B. Ahmed, K. N. Tahsin, S. Khan, M. Hoq and D. Alam, "Design and implementation of a microcontroller based flammable gas detector and automatic alarm system to ensure the industrial and domestic safety", Asian Journal of Science and Technology, vol. 9, pp. 7425-7429, January 2018.

[6]. N. B. Ahmed, S. Khan, K. N. Tahsin, F. Hafiz, M. Hoq, M. A. S. Haque and F. Akter, "A low-cost microprocessor based line-follower robot for monitoring temperature and humidity under adverse conditions", International Journal of Pure and Applied Mathematics, vol.119, pp. 12611-12619, Academic Publications Ltd, May 2018.

[7]. N. B. Ahmed, S. Khan and F. Hafiz, "Design and implementation of an autonomous wireless information transceiver robot", Proceedings of National Conference on Electronics and ICT, Bangladesh Electronics Society, Dhaka, April 2017

[8]. Javed P., O'Kelly, K. Dick, M. Lucey, "Wireless technology evolution and impact on the access network", Proceedings of IEEE International Conference on Personal Wireless Communications, IEEE Press, New York, 1994.

[9]. H. T. Yew, E. Supriyanto, M. H. Satria, Y. W. Hau, "Adaptive network selection mechanism for telecardiology system in developing countries", IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), IEEE Press, New York, 2016.

[10]. S. Khan and I. A. Mamoon, "Telemedicine services at the Diabetic Association of Bangladesh with technical collaboration of Grameen Phone", Prospects and Problems of Mobile and Land Phones in Bangladesh, 2005, pp.88-95.

[11]. S. Khan and I. A. Mamoon, "Implementation of telemedicine in Bangladesh, with Grameenphone Healthline 789 project as case study", Proceedings of National Conference on Electronics, Information, and Telecommunication, Bangladesh Electronics Society, Rajshahi, Bangladesh, 2007.

[12]. S. Khan, Telecommunication Engineering, S. Khan, June 2016, Dhaka, Bangladesh.

[13]. S. Khan and I. A. Mamoon, "Implementation of a GSM based nationwide medical call center", Journal of Telecommunications UK, vol. 8, pp. 51-55, 2011.

[14]. S. Khan and I. A. Mamoon, "An implementation of a GSM based nationwide medical call center", Proceedings of International Conference on Electronics, Computers, and Communications, Bangladesh Electronics Society, Rajshahi, Bangladesh, 2008.

[15]. S. Khan, I. A. Mamoon, and M. A. Jabbar, "Telemedicine in Bangladesh with Diabetic Association, Bangladesh- Grameen Telecom project as case study", Proceedings of International Conference on Electronics, Information, and Telecommunication, Bangladesh Electronics Society, Rajshahi, Bangladesh, 2005.

[16]. S. Khan, I. A. Mamoon, and M. A. Jabbar, "Telemedicine in Bangladesh with Diabetic Association, Bangladesh- Grameen Telecom project as case study", Proceedings of National Conference on Electronics, Information, and Telecommunication, Bangladesh Electronics Society, Rajshahi, Bangladesh, 2007.

[17]. S. Khan, Power System Operation and Protection, Second Edition, S. Khan, February 2020, Dhaka, Bangladesh.

[18]. S. Kumar, N. S. Gayathri, D. Nivedhitha and A. S. Priyanka, "A cost-effective Arduino module for bedridden patient's Respiratory Monitor and Control", International Journal of Advanced Research Trends in Engineering and Technology, vol. 2, 2015.

[19]. The Statistics Portal, "Number of smartphone users worldwide from 2014 to 2020 (in Billion)", online publication.

[20]. T. S. Arulananth and B. Shilpa, "Fingertip based heartbeat monitoring system using Embedded Systems", Proceedings of International Conference of Electronics, Communication and Aerospace Technology, IEEE Press, New York, pp. 227-230, 2017

[21]. M. S. Mahmud, H. Wang, A.M.E. Alam and H. Fang, "A wireless health monitoring system using Mobile Phone accessories", IEEE Internet of Things Journal, vol. 4, pp. 1-10, IEEE Press, New York, 2017.

[22]. R. Schoen, "Women and rural industrialization: Garment production reaches old land and new labor in Bangladesh", Elsevier Journal of Woman's Studies International Forum, vol. 75, 102248, 2019. Doi: 10.1016/j.wsif.2019.102248

[23]. C. Holland, A. Rammohan, "Rural women's empowerment and children's food and nutrition security in Bangladesh", Elsevier Journal of World Development, vol. 124, 104648, 2019.

Large Scale Distributed Data Processing for a Network of Humanoid Telepresence Robots

Irvin Steve Cardenas
Advanced Telerobotics Research Lab
Computer Science
Kent State University
Kent, Ohio, USA

Pradeep Kumar Paladugula
Advanced Telerobotics Research Lab
Computer Science
Kent State University
Kent, Ohio, USA

Jong-Hoon Kim †
Advanced Telerobotics Research Lab
Computer Science
Kent State University
Kent, Ohio, USA

***Abstract*—We present an open-source data lake architecture implemented to store and process data from robotic systems at large scale. In particular, we leverage our architecture for the use case of processing data from a network of humanoid telepresence robotic avatars that are controlled by human operators wearing immersive telepresence control suits. Our architecture leverages well-established open-source technologies and integrates into existing robot frameworks and middleware such as Robot Operating System (ROS) and Data Distribution Service (DDS).**

***Index Terms*—cloud robotics, big data, IoT, distributed systems**

I. Introduction

The present and foreseen human migration from the physical (on-site) workplace to the (remote) home office, is a key indicator that technologies related to teleconferencing and telepresence will continue to gain traction over the years. At the office, this is important as teams have to adjust their dynamics to incorporate a mixture of members working on-site and remotely. Video conferencing has become the number one choice to facilitate such interactions. There is also a number of "telepresence" video conferencing robots on the market – or, as we call them – "iPads on a stick with wheels". Although, the latter two are the most common solutions, what is not apparent at first glance is that video conferencing - either through a camera or an iPad on a stick - does not address one key factor that brings together remote and physically-present individuals.

This key factor is what telepresence research refers to as the feeling of "being there". Humans, by nature, incorporate both verbal communication and non-verbal cues such as hand gestures. Furthermore, we have to consider that interaction between humans do not only take place in meeting rooms or at someone's desk. Humans are always in motion. In other scenarios, besides office teleconferencing, the need to feel that someone is present or to extend an intimate gesture is often desired. For example, when videoconferencing with loved ones, one may wish to extend a hug, a handshake, or caress one's child to sleep. Videoconferencing does not afford such interaction. In other cases, one may want to remotely explore a foreign tourist location, prior to a physical visit. Conversely, one may be constrained to make such physical visit due to various reasons, such as monetary limitations or physical disability. In present times, a pandemic such as the one we face requires individuals to limit their interactions with the outside world and has reduced the number of activities which individuals may participate in[1].

In either case, being able to feel present - to see, to hear, and even feel the remote environment would be more desirable than a mere video call or a live video tour.

Our work relates to developing immersive humanoid telepresence robotic avatars, pictured in fig. 1 (a) and (d), where the operators are afforded the sense of being present by leveraging control interfaces such as the one in fig. 1 (b) and (c). The latter garment: (1) allows the operator to receive tactile and force feedback from the remote environment, (2) operate the robots via natural motions, and (3) monitors the physiological state of the operator. In this paper, we particular focus on the development of a distributed data processing system for a network of immersive telepresence robotic avatars – where high amounts of data is bi-directionally streamed and captured. This data includes but is not limited to: motion control commands via motion capture wearables and direct teleoperation controllers, tactile and mechanical forces experienced by a robot, video and laser data from RGB and/or depth-sensing laser cameras, network telemetry, system state information, and operator state information. We elaborate on the data properties in section III.

To this end, we seek to answer the following: "how can we process such large and diverse amount of data, when a network extends to thousands or hundreds of thousands of human-robot pairs?" and most importantly, "what value can we extract from such data and for what purposes?"

Overall, the proliferation of robotic systems in research, commercial or personal environments, will become part of the Internet of Things (IoT). Inherently, with such interconnectivity of devices and end-systems, the production and consumption of data by such is bound to increase. From the perspective of robotic systems, this brings emergent opportunities to leverage data analytics not only for the sole purpose of (1) improving the performance of a robotic system, but also for (2) improving human-robot interaction, and (3) understanding the ecosystems in which such robots are deployed. In our use

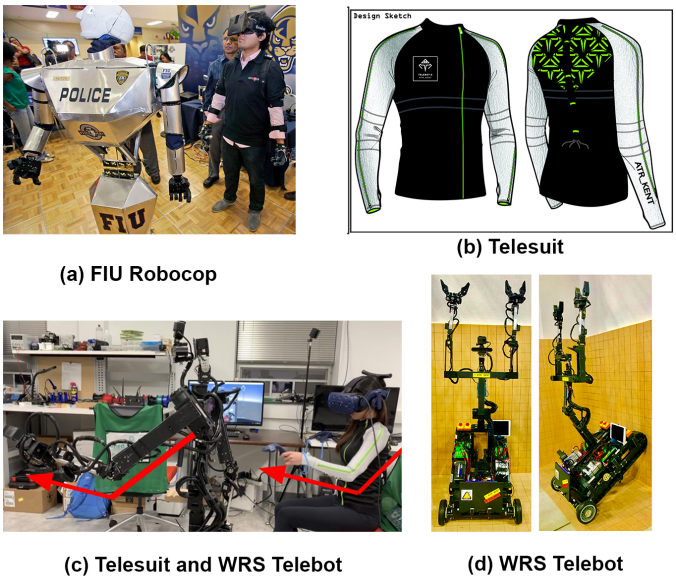
(†) The corresponding author's email: jkim72@kent.edu

[1]For example, as of June 2020, there is still a limited number of international flights, and various tourists attractions still remain close.

case of immersive humanoid telepresence robots, our goal is to create a virtuous cycle where we can leverage data collected from the robot (and the remote environment) to enhance the operational experience, and vice-versa.

In this paper, we present the design and implementation of a data lake architecture developed to process and analyze data produced and captured by immersive humanoid robotic avatars and their operators. Unlike, existing literature that presents a high-level view of big data systems in robotics (e.g. cloud robotics), we provide a thorough look into how to integrate such technologies and, most importantly, we open-source details of our implementation for researchers to explore and deploy. This implementation, makes little to no use of proprietary solutions or blackbox technologies [2] which minimizes vendor lock-in and allows to freely experiment..

Section II, presents the previous work related to cloud robotics. Section III, presents an overview of the data we process, challenges, design goals. Section IV, presents our architectural considerations and components.Section V presents our system evaluation. Finally, section VI presents our future work and section VIII concludes our paper.



(a) FIU Robocop

(b) Telesuit

(c) Telesuit and WRS Telebot

(d) WRS Telebot

Fig. 1. Telepresence Robots and Telesuit

II. State of the Art

The overall concept of a data lake represent a large, centralized, repository that stores raw data of diverse form [13], [16]. Diversity refers to data that has either a (1) structure (e.g. in proper row/column format), (2) unstructured data that has "no shape" (e.g. text, video, audio), or (3) semi-structured data - a mix of both (e.g. data that has a mixture of tags in JSON or XML format). The notion of a "data lake" can be visually understood by portraying data as water which may flow from different sources and which must be captured somewhere to extract value. This is contrary, to bespoke flows of data that arrive at different silos (data silos). This would force individuals to visit each silo to query or analyze data, and does not afford an overarching view of all available data. The concept of a data lake, proposes that all flows of data be capture on a single pseudo-location [3] (a lake), and processed through a scalable data pipeline which allows for rigorous control of data (what kind of data, how it is accessed, and who accesses it). This pipeline also creates opportunities to automate and scale machine learning / artificial intelligence pipelines, as we later discuss.

The concept of a data lake, sprung from the ever increasing popularity of "the cloud" and distributed compute/storage technologies, and the need to apply this technology to make sense of large amounts of data. The includes distributed technologies such as that within the Hadoop ecosystem [2] or Spark [34] which allow for efficient storage and processing of big data [11]. "The cloud", simplified the management of physical hardware required to run such distributed storage and compute software, by allowing developers to offload the management and maintenance of such hardware to big corporate giants such as Amazon, Google and Microsoft.

With the advantages of the cloud, the field of Cloud Robotics [20], [31] emerged in an effort to scale computationally expensive algorithms by moving computations to the cloud and proposing methods for robotic systems to share large amounts of data hosted. E.g. visual SLAM approaches that leverage cloud computing were proposed [5] that allow for large number of landmarks to be processed, and even web tools for cloud robotic systems [28] have been presented. This also sprung the business model of Robot-as-a-Service (RaaS) [9]. Projects like Robot-Cloud, Brass and XBotCloud also emerged during the early years of Cloud Robotics. In 2012, Robot-Cloud [12] proposed leveraging the Map / Reduce framework [4] to offload expensive computations, a system for storage, and robotic services such as object recognition, path planning, localization etc - to the cloud. More recent work includes BRASS (Brekeley RAaaS Software) and XBotCloud. BRASS offered Dex-Net - a recommender system for grasp planning [27]. XBotCloud [23] provides a ready to use software through web-interface. This includes object recognition, object segmentation, pose estimation, trajectory generation, etc. Although, these projects leverage the concept of the cloud, their application has been limited to in-lab research environments, and data processing as well as the management of data captured by such robots is not addressed in such literature. By implementation, multi-sensory robotic systems, such as humanoid robots or autonomous vehicles, in-the-wild will capture and consume data that is not limited to imagery

[2]Cloud-as-a-Service technologies, such as those offered by Amazon AWS, Google Cloud, or Microsoft, are not open.

[3]Data lakes leverage distributed compute and storage technology, but conceptually present a single unifying interface through which the data may be stored and accessed.

[4]Map / Reduce is computational framework of Hadoop where the data is divided into multiple chunks and sent to different servers where data is processed and the solution is added on the received server machine from the received processed information.

related to path planning or grasping algorithms - as focused on in literature. Other types of data may be collected such as sensory feedback from unconventional control interfaces – e,g. figure 1 (b), LiDAR, or time-dependent data such as heart pulse rates, and even audio signals.

Two projects, in particular, that closely relate to the application of a data lake in the field of robotics include: RoboEarth - dubbed as the "world wide web for robots" or a "common brain". [30], and DAvinCi - an abbreviation of Distributed Agents with Collective Intelligence [3]. During its lifespan, RoboEarth focused on providing a knowledge sharing database for autonomous robots to perform tasks such as object detection. The overall goal was that by querying the RoboEarth system, robots can improve their performance speed and reliability since known models or knowledge of different environments could be shared and then enhanced by the robot - rather than requiring the robot to collect new information. The database stored other properties such as object types, location where the object was detected etc. The integral components of the database are hierarchical tables, a graph database which contains the data and their semantic relationships. The queries are passed through semantic querying language. This project was funded by the European Union from 2010 - 2014, and is no longer active. Similarly, DAvinCi focused on developing a platform for offloading computational intensive tasks. They implemented this system by combining Hadoop's Map/Reduce framework [2] with the Robot Operating System (ROS) architecture [3]. Unfortunately, the implementation details of this project have been kept private - which we note is common in other cloud robotics literature.

Although such nascent efforts are valid, various architectural and implementation considerations were overlooked. Overall, during the early days of Cloud Robotics there was a keen focus on the actual implementation of systems that leveraged the cloud for any purpose, and less considerations for latency and bandwidth constraints. Currently, the emergence of Fog Robotics [8] aims to tackle such issues through the integration of a fog computing architecture into robotics - promising lower levels of transmission latency. In terms of architecture, leveraging multiple hierarchical tables for object models (as RoboEarth) can provide simplicity, but is not scalable once we consider a high number of many-to-many relationships that may need to be modeled. Overall, different databases or storage solutions have their own advantages and tradeoffs, which must be considered when developing a data system. In our work we assess the data produced by the robots and operators and leverage adequate technologies that benefit our storage and analytic goals.

Additionally, systems presented in cloud robotics literature do not consider the evolution of the data [5] and assumes that data schemas are known before hand or that a robotic system may not evolve to include additional sensors or functionality. As a side note - beyond cloud robotics, there is a lack of consideration for schema evolution still present in research software like ROS Bag [14], which assumes that the data produced by a robot may not change over time and hence not allowing messages with progressively different schema to be tied to the same topic. But, all in all, we must consider that earlier efforts in Cloud Robotics such as DAVinCi were not exposed to technologies beyond the Hadoop stack. For example, in recent years Spark [34] has become a key technology in the world of cloud computing - it offers the possibility of real-time and near-real time processing [6]. With such modern technologies, we can develop more robust cloud-based robotic systems, and even extend such systems to leverage the fog computing paradigm [8], [25].

More recent efforts to simplify the use of "the cloud" in robotics has been made by industry giants such as Amazon and Intel. With Intel developing it's Insight Platform with a key focus on geospatial visualization or geovisualization intelligence; as well as focused on IoT [17]. Though, their vision looks to address sensory information and real-time analytics; the architecture presented is very high level and presents a broad perspective on a data infrastructure for sensors, drones and other robotic system. Similarly, Amazon recently released RoboMaker [4], a platform that hosts ROS on in the cloud - similar to what was experimented in the DAvinCi project.

The concept of a `data lake` can help us grasp what it is like to handle large amounts of data from a network of robotic systems and IoT devices. Robots may collect different types of data: unstructured (image, video, lazer), structured, time-dependent / eventful data such as sensory data, or a combination of both (semi-structured information). The latter brings into play the notion of variety in data. Data produced by a single robot or by various robots in-the-wild may be of varying size: from a few megabytes per second, to a gigabytes or terrabytes per day [33] [15]. This data will also be produced at different rates and may need varying processing reates - e.g. real-time analytics, batch processing, after-the-fact analysis.

In this work, we focus on applying the concept of a data lake to ingesting, processing and analyzing data produced by a network of (1) immersive, (2) humanoid, (3) telepresence robots. We number these words as they define the key requirements and constraints to building such data lake. We can trace back the concept of telepresence to Marvin Minsky, who described the goal of telepresence as developing robots that give the operators the feeling of "being there" [22]. That is, developing robotic avatars that allow human operators to be presence in a remote environment. This vision is, of course, contrary to the contemporary consumer telepresence robots available on the market - which can be regarded as mobile video conferencing platforms on wheels [29] [19]. To achieve the feeling of "being there" we need technologies that allow the operator to not only actuate (i.e. teleoperate) a remote avatar, but also technologies that allow the operator to see, hear, and feel the sensations of the remote environment - overall, technologies that allow the

[5]Robot data may change over time. That is, their schema is bound to evolve and we must account for this evolution in any data-related system.

[6]Near-real time processing is where information is processed on basis of micro-batches. Real-real time processing is individual pieces of information, which are processed as the data is sent [13]

operator to feel what the robot is feeling. Such technologies include humanoid robots such as the ones in fig 1 (a, d) which are equipped with the tactile sensors, and developing wearable interfaces such as the one in fig. 1 (b, c) which allows the operator to control a humanoid robot through natural arm motions, and which allow an operator to experience remote tactile or mechanical sensations.

III. Preliminaries

A. Data-driven Interactions

In our research, the overarching goal of processing and analyzing data from a human-robot pair is to create a virtuous cycle where we can leverage real-time and after-the-fact analysis towards improving the interaction between the human and the humanoid telepresence robotic avatar. For example, capturing real-time motion control data over a time-period can allow us to develop manipulation models that can infer future motion control commands of an operator. This, can allow the control of a robot to become more fault-tolerant to bandwidth-constrained environments, or message drop-outs, which can reduce the number of motion control commands a robot receives. Other use cases involve the real-time collection and processing of video information to enhance awareness of a robot's environment. In parallel, data representing the operator's physiological state can be leveraged to infer the overall state of the operator when controlling the robot. The latter, for example, can lead us to identify interactions or control gestures that may cause physical or metal discomfort which often seen during direct teleperation of robots, and in immersive telepresence control [7]. Furthermore, the latter information can assist in defining user profiles that indicate which control gestures or interactions the robot may need to automate or semi-automate to improve the overall human-robot interaction. The latter relates to our work on assisted autonomy and kinesthetic learning algorithms.

B. Data Overview

A sample aggregate view of the data produced and consumed by an operator utilizing the Telesuit is presented in listing 1. The data can be categorized as follows: (1) control commands, (2) immersive feedback, (3) operator physiological state. This data is sent and received through a direct socket connection between the robot and operator.

`Control command` data is produced by the motion capture sensors located across the control garment. These sensors model the operator's torso as a kinematic chain with rotations and angles of each joint. The latter can then be used to control the corresponding joints of the humanoid robot. Production of this data varies depending on the hardware / system architecture. `Immersive feedback` data is tactile and force feedback data sent from the remote location, where the robot is located, to the operator's control garment. The data incites vibrotactile motors located across the suit that represent the regions of the robot where force was experienced. `Operator physiological` data, is data captured by the Telesuit. This includes ECG data, Galvanic Skin Response (GSR), and body temperature.

```
{"ocu_id": "5ed2c0243d43a4351dd28c4c",
  {"time": 4.016,}
  {"timer": "00:00:04",}
  {"started_at": "2020-05-30T20:20:52.961Z"},
  {"joint_l_arm": [88.00, 45.00, 81.00, ...],
  {"joint_r_arm": [42.00, 81.00, 33.00, ...],
  {"heart_bpm": 89},
  {"p_suit": "3.94"},
  {"t_sub": "31.8"},
  {"v_fan": "39977"},
  {"p_o2": "776.05"},
  {"rate_o2": "0.9"},
  {"battery_in": 99},
  ...
}
```

Listing 1. Sample Operator Data

Listing 2, shows a condesend aggregate view of the numeric and qualitative data produced and captured by a telepresence robot. Particularly, individual motor control feedback (i.e. encoder data) is sampled at rate of 125kHz. Motor feedback is sample at 30 Hz. Furthermore, tactile and mechanical pressure information is captured when forces above an arbitrary threshold are produced. This indicates that the environment or another agent is interacting with the robot's synthetic skin. This is data is continuously produced and cached prior to being stored. Additionally, data from a pair of RGB cameras, depth sensing information, and cloud point data is captured by the robot. LiDAR samples 1.5 million data points per second

```
{"id": "5ed2c0243d43a4351dd28c4c",
  "time": 4.016,
  "timer": "00:00:04",
  "lat_lon": [ 40.758896, -73.985130 ],
  "started_at": "2020-05-30T20:20:52.961Z",
  "l_arm_pos": [88.00, 45.00, 81.00, ...],
  "r_arm_pos": [42.00, 81.00, 33.00, ...],
  "l_arm_temp": [65.00, 65.00, 70.00, ...],
  "r_arm_temp": [65.00, 67.00, 65.00, ...],
  "p_suit": "3.94",
  "t_sub": "31.8",
  "v_fan_front": "39100",
  "v_fan_rear": "38888",
  "battery_out": 90,
  "batt_prim": 90,
  "batt_sec": 99
  ...
}
```

Listing 2. Sample Robot Data

C. System Design Goals and Challenge

Our implementation optimizes for availability in message publishing and does not focus on data loss. This is due to the fact that hardware and internal software components of the robot and operator control unit can produce variant sub-second data containing duplicate state and event information.

Overall, our focus is on being able to process, real-time, batched and streams of data from hundred to thousands of robot. Data captured by each robot is between 1 to 5 terrabytes per hour depending on the LiDAR configuration, and it is produced at varying rate. The operator control unit, which includes the Telesuit and software produces could produce between 5 to 20 GB of data per hour. Although out-of-the-box or blackbox enterprise solutions exist in the market, such as AWS Data Lake, our purpose for implementing this open-source system is to leverage this system as a modular research platform where additional components and features (e.g. encryption, blockchain-integration) can be integrated without vendor lock-in and dealing with additional expenditures.

IV. System Architecture

A high level representation of our system is presented in Fig. 2, and a data flow diagram is presented in figure 4. We present the architecture as two environments: (1) local / remote environment - software components that are part of the operator control unit / robot respectively, briefly discussed in section IV-A, and (2) cloud environment - data lake software components, discussed in sections IV-B, IV-C and IV-D.

A. Peer-to-Peer Communication and Local Data Sinks

Direct teleoperation and feedback data is sent from the operator to the robot, and vice-versa, through a secure socket connection. Internally, the operator control unit and the robot's internal software implements a fan-out messaging strategy with *n* number of channels. One channel in particular delivers real-time data to a p2p socket communication broker – the broker manages the delivery of the data from the operator to the robot or vice-versa. The other internal messaging channels stream the socket data, and other system data, into an *n* number of `data sinks`. A data sink acts as a buffer used to cache data until it can be flushed to the data lake – this buffering mechanism avoids overwhelming the data lake's proxy service layer with continuous sub-second requests and also optimizes for Kafka's ideal message size of around 1 KB. We can understand the role of a data sink's buffer as a temporary append-only log which gets appended to a global log. This analogy of a log is similar that of Kafka's internal log.

Sinks are created for different event and state `data streams` produced by the system's hardware or internal software processing. For example, the telepresence robot in fig. 1 (d) makes use of two hardware motor controllers. Each motor controller controls the motors in the left and right arms respectively, and publishes motor feedback (i.e. motor temperature, position, amperage). This data is piped to a "motor joint state" data sink. Similarly, in the robot's side for example, one data sink collects robot state information which includes battery life, external temperature, and geographical location.

Within the operator control unit, DDS [24] is used as a messaging middleware with a software wrapper that subscribes to ROS and DDS topics. DDS is leveraged because it's lightweight, language agnostic, and provides quality-of-service management - configurable parameters that control the behavior of a DDS system. These parameters include: resource consumption, fault tolerance, and communication reliability. The latter is also useful towards simulating network degradation during our simulations. As a middleware, DDS enables redundant data sinks to exist outside of the central computer that handles the video / audio streams, and communication with the Telesuit - thus making our system more resilient to failure at the data lake's proxy service layer.

Similarly, data sinks are used in the robot software architecture for temporary storage of numeric and video data. ROS Bag is used to collect data internally. The RGB camera video stream is multiplexed using a Video4Linux kernel module that allows the robot to create multiple virtual video devices. One device is used to stream video to the operator, and another virtual device's video feed is recorded and placed in a video data sink for later stored in the data lake.

B. Data Ingestion

There are two goals for ingesting data into the lake, the first is for real-time or close to real-time analytics, and secondly for after-the-fact processing / data refinement. The stakeholders of real-time analytics are not just individual robot operators, but also network administrators who can monitor the state of the network of robots and get a holistic view of properties such as up-time of robots, network latency, and physical traffic hot spots. This information can then be used to enhance the human-robot interaction or overall experience of controlling a robot. E.g. operators can be alerted of communication dead zones, or even visual anomalies that are captured by other robots. This is further discussed in section IV-C.

Ingestion begins through a Proxy Service Layer where a REST service is deployed on Nginx web server that acts as a load balancer. Each REST endpoint maps to a topic. This can create a bottleneck and throttle the transfer of data between Kafka broker and the consumer, but also removes the need for maintaining and rotating SSL certificates, and dealing with encryption / decryption on the Kafka servers. This approach models the service offered by Confluent [10].

The data from each respective REST endpoint is then used by a Kafka producer to publish using a topic name. A load balancer is necessary as the endpoints may be overwhelmed by an increasing number of operator-robot pairs. Within each message an operator's unique identifier is used to generate a partition key in Kafka. Partition keys make sure the information of a specific user is not spread across different partitions but rather sent to a single partition. Similarly, a robot's unique identifier is used to generate a partition key.

Overall, deciding on the number of partitions in Kafka is considered a "million dollar question". It is very hard to get it right without iterative testing. However, we employed the following guidelines, based on the ample literature on Kafka, to decide on the partition configurations:

1) The least number of partitions for a given topic, should be based on the maximum consumers during peak time

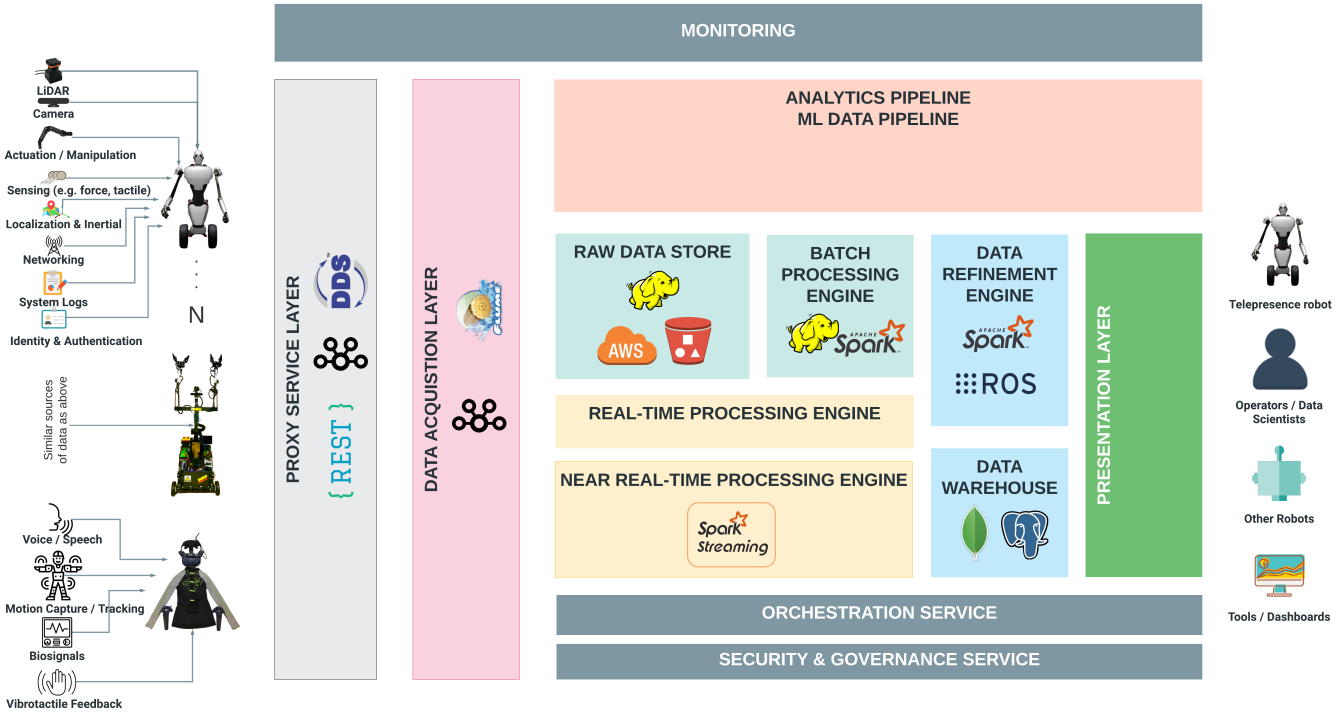


Fig. 2. High level view of our architecture

2) Once the number of partitions is changed, the data ordering is not guaranteed
3) Arbitrary partition configurations may be useless
4) Replication factor should **not** be compromised. Replication factor should at a maximum be 4, usually 3
5) No more than 2,000 to 4,000 partitions per broker
6) A cluster should have a maximum of 20,000 partitions. Otherwise, a multi-cluster architecture can be used
7) Unless redundancy is necessary, the number of brokers should not be greater than the number of partitions. Otherwise, additional brokers will become idle

Although, Kafka producers can publish to different topics, for performance we isolate one producer for each topic. Once the data has been published through Kafka, different consumers can subscribe and read the data. The load balancer is configured so that a new Kafka producer is initiated when the REST servers scale.

C. Storage, After-the-fact and Real-time Processing

Although, a partition is the actual unit of storage in Kafka (i.e. what is physically stored on the disk), and log `retention` can be set to ”forever” – Kafka is particularly used as a scalable message queue. With that consideration, consumers of data published through Kafka can push the messages to queriable data storage systems or unstructured file systems that can allow processing at a later time. Additionally, Kafka provides library extension such as the Streams API which allows processing and analyzing data in Kafka. This does not involve querying Kafka directly, but rather generating a data view on the data so that a stream processing application can run low latency queries against it.

For persistent storage we leverage the Hadoop File System (HDFS) and MongoDB. Apache Flume is used as a `stream processor` that subscribes to a topic and places the topic messages into a given directory in HDFS. MongoDB is used to store data of operator-robot control sessions. After each session, an analytics process can directly query MongoDB, this contrary to storing data in HDFS which requires additional libraries and tools to structure and query the data. On the other hand, HDF is used for long-term storage of the data in its unstructured form. In the enterprise world, HDFS is similar to AWS S3 Bucket.

Real-time processing is performed leveraging the Kafka Streams API and Spark Streaming. One pertinent use case was discussed at the beginning of section IV-B. This use case relates to processing network information such as robot uptime, network communication latency, to alert robot operators of any issues they may face during control. With Spark Streaming we can run analytics such as anomaly detection and network monitoring etc.

D. Analytics and Machine Learning

Fig. 5 presents our implementation of a real-time machine learning pipeline based on [18]. Presently, we focus on generating real-time analytics on communication network data, operator physiological data, and after-the-fact analytics on video data that is stored in batches in the lake. Predictions are done on edge computing device rather than running predictive

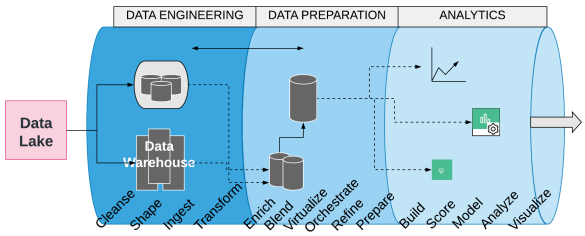


Fig. 3. Processing Pipeline

running algorithms on cloud premise, one advantage of doing this is guaranteed low latent predictions.

Real-time anomaly detection within the network communication data is performed using a Spark Stream Processor that filters messages by city and then runs anomaly detection, the algorithm is trained before hand, however, the model should be updated for the incoming data, following are few scenarios:

- If new predictions have high prediction error
- A scheduled model update during late nights, e.g 12AM when low traffic is expected
- Updating the model for continuous in flow of data

Choosing an anomaly detection model is really difficult as algorithms tend to over-fit on training samples, so choosing an predictive algorithm is really crucial to this matter. Unlike, network data video is not stored or processed in real-time. Fifteen to thirty minute RGB video streams from the robot are first cached in the robot and the stored in the lake's HDFS in the following directory structure: `robot_uid/ocu_uid/date/sess_id/batch_num`. These video files are then extracted from HDFS using a HttpResquest on to WebHDFS port which allows for queries based on a robot's ID or operator control unity ID. The video data is then fed to a Tensorflow video recognition model.

Overall, there are two verticals of predictive analytics for our architecture both leverage the Kafka message queue and use technologies such as Kafka Streaming and Pyspark. Both of these address real-time and near-real time analytics. For example, with the help of `pyspark.streaming.kafka` [26] we are able to subscribed to certain topics to read a stream of data. Another vertical of our analytic pipeline finds insights from historical data, we leverage PySpark for this. Presently, our use of the JSON format makes it easy to manipulate data as it's compatible with Spark's dataframes, where further computation is done.

There are various use cases in which we can leverage analytics to improve the human-robot interaction and the overall experience of controlling a robot. We can also imagine future scenarios where analytics could play a role. For example, in a shared economy, similar to ride sharing applications – we can imagine that humanoid telepresence robots would be available for rent through a RaaS model. Consequently, these robots could leverage a dynamic pricing model that would charge human operators prices according to robot demand, weather, holiday events, or even expected lease time.

E. Visualization

Visualizations are performed using a Tableau server, connected to an existing data warehouse. Two types of data warehouses are supported - MongoDB and PostgressSQL, both of which have data Connectors available by Tableau. A data warehouse is the end result of extracting data from the lake, transforming it and loading it. This transformation could include restructuring the data or cleaning it. Our current implementation does not validate any data through a schema.

For example, motion control data from the Telesuit, collected during a telepresence session, is simply extracted from HDFS and loaded in PostgressSQL as row-column values with the primary key being a combination of the time and operator's unique ID. Motion control data is simply, well-structured, time series data. Migration data into a warehouse is done through software scripts that copy the data from the respective HDFS directories and load it into a warehouse. Future implementation, aims to automate this process and make it on demand; optionally, the warehouse could be continuously populated by using a Kafka Connector (e.g. Kafka MongoDB Connector).

Overall, a data warehouse should meet the requirements of the data and business requirements. For example, within the ROS framework, ROS Bag is used for collecting arbitrary data. This includes cloud point data, which could amass gigabytes of data in minutes. E.g. one LiDAR with 2 RGB cameras generate data at a rate of 2 GB / minute on average. ROS Bag is a store of serialized ROS messages of any kind. Internally these messages are stored as indexed timestamped chunks of data - i.e. a structured timeseries. When storing LiDAR cloud point it may be more effective use an unstructured database, as explored by [6], [32]. Hence, a MongoDB warehouse would be better suited for loading cloud point data.

F. Service Orchestration, Security and Governance

Presently data extraction from HDFS or MongoDB, and migration to a data warehouse is performed through manual execution of scripts. Future work, implements automatic execution of the extraction scripts using managed CRON jobs. Current implementation only uses SSL certificates in the Service Proxy Layer, and OAuth to access the data for visualization. Future / ongoing work explores data governance techniques that introduce PKI encryption and immutable access logging through blockchain.

V. System Evaluation

A preliminary test was performed using the simulated production and consumption of data by 100 operator - robot pairs. A 3-broker Kafka cluster was used, running on Intel Xeon processors (up to 3.3 GHz) with 64 GB RAM. The same specs were used for MongoDB (recommended 4 GB of RAM) and Hadoop (minimum recommended requirement of 8 GB). For a single robot-pair interaction, publishing messages of size of 1 MB resulted in Kafka producer rate of 8.91 ms for production of data, and 13.53 ms for consumption of data.

We note that our system configuration and evaluation is meant to validate the functionality and ideal throughput of

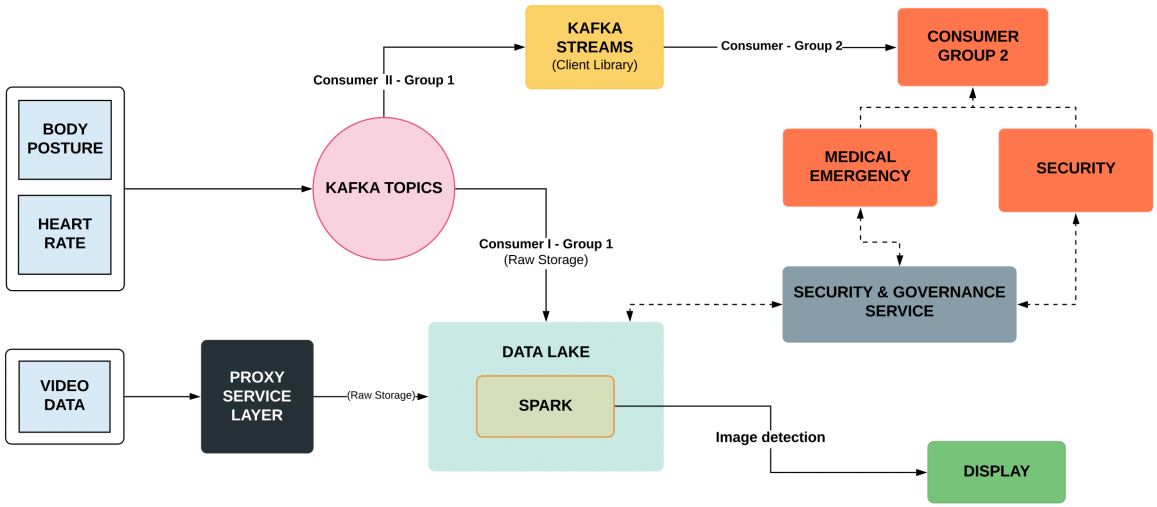


Fig. 4. Data and Analytical Flow Diagram

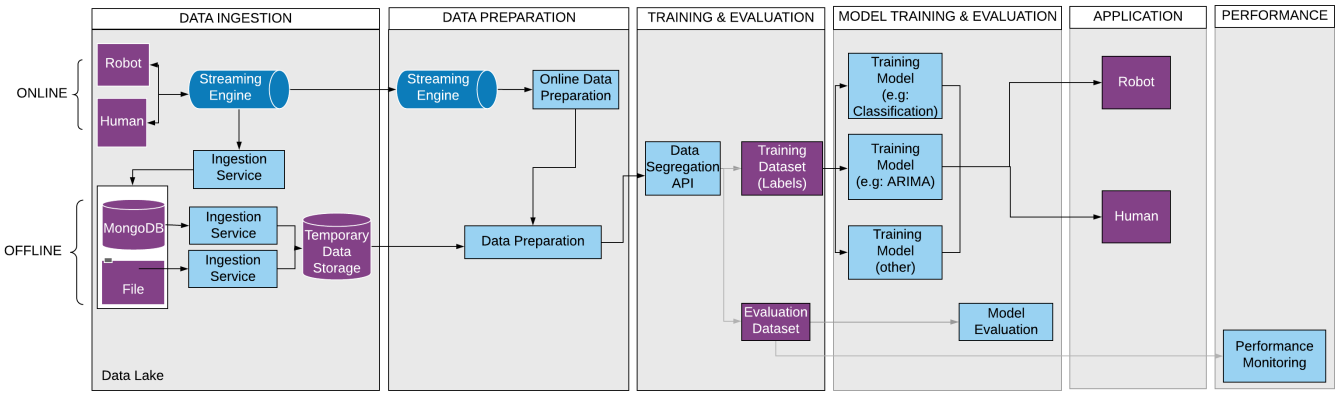


Fig. 5. Machine Learning Data Pipeline

our system architecture. Our present evaluation configurations differ from the final large scale deployment - for large scale please refer to section IV-B. In our implementation we have considered topics for operator and robot with three replication factor, three partitions and three brokers.

On the other hand, extraction and loading time of data from HDFS to a MongoDB is linearly correlated to the size of data. Internally, the data extractor process simply copies the data found in a given HDFS directory (no additional overhead). Then, the data loader transforms the data (if needed) and loads it into the MongoDB warehouse.

Since video is only used for after-the-fact processing it is arbitrarily chunked locally by the robot on 15 minute intervals. Further evaluation on latency and throughput is necessary to adjust chunking intervals and also video resolution.

Architecturally, the requirements for a `Proxy Service` can create a bottle neck for continuous or real-time storage of data. Each data sink flushes the data once it has reached an arbitrary size of 1MB or if more than a given threshold time has take place. This arbitrary size is larger than that recommended by existing benchmarks, that identify that the best performance for Kafka occurs with 1 KB messages. Larger messages such as 10 MB to 100 MB can decrease throughput and significantly impact operations.

VI. Future Work

Our future work considers data privacy measurements such as encryption, differential privacy, and the use of blind signature schemes [1] to preserve the anonymity of robots and operators that are sharing telemetry data. We also consider data compression and evolving data schemas - presently the data produced by the operator control unit and the robot is transformed into JSON prior to being sent to the data lake. From an implementation perspective, we are exploring different scalable technologies that allow us to temporally store ROS topic data, given that ROS Bag presents some storage limitations and the ROS 1 architecture adds unnecessary processing overhead. Additionally, we explore a means

to decentralized the storage layer by leveraging blockchain-based decentralized storage solutions, and also delve into the application of smart contracts and cryptoeconomics as a means to allow operators and robots to monetize data captured during telepresence session.

VII. Limitations

In our architectural design we have considered several factors which are appealing for real-time applications such as (1) availability, (2) parallelism and (3) high throughput. However, there are performance limitations (in addition to what was mentioned in section IV-B) to configuring a Kafka cluster. Different configurations bring a trade off between the latter three factors. The following are a few limitations we expect to encounter on a large scale implementation:

Number of Kafka Producers: Different topics have different throughput of messages. That is, we need to consider that not all telepresence robots have the same number of sensors / produce the same data. Hence, further analysis would be necessary to establish whether its efficient for a producer to produce messages to more than one topic. **Number of Kafka Consumers:** An increasing number of consumers would require additional load balancing from Zookeeper, and would disrupt the number of required partitions. **End-to-end Latency:** This is based on various factors such as partitions, message acknowledgment, and network conditions.

VIII. Conclusion

This paper presents the end-to-end architecture and implementation of a data lake system for storage and processing of data from a network of immersive humanoid telepresence robotic avatars. The overall goal of this system is to facilitate the real-time and after-the-fact extraction of knowledge and insights that could improve the human-robot interaction that takes place when humans wearing immersive control garments operate remote humanoid telepresence robots. This architecture leverages fully open-source technologies. Further analysis and the open sourced implementation is available in [21].

References

[1] K. M. R. Alam, S. Mahmud, and M. N. A. Khan. A comparison between traceable and untraceable blind signature schemes through simulation. In *2013 International Conference on Informatics, Electronics and Vision (ICIEV)*, pages 1–4, 2013.

[2] Apache. Hadoop. https://hadoop.apache.org/.

[3] R. Arumugam, V. R. Enti, L. Bingbing, W. Xiaojun, K. Baskaran, F. F. Kong, A. S. Kumar, K. D. Meng, and G. W. Kit. Davinci: A cloud computing framework for service robots. In *2010 IEEE International Conf. on Robotics and Automation*, May 2010.

[4] AWS. RoboMaker. https://aws.amazon.com/robomaker/.

[5] K. Ayush and N. K. Agarwal. Real time visual slam using cloud computing. In *2013 Fourth International Conf. on Computing, Communications and Networking Technologies (ICCCNT)*, pages 1–7, July 2013.

[6] J. Boehm and K. Liu. Nosql for storage and retrieval of large lidar data collections. *ISPRS - Intl. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-3/W3:577–582, 08 2015.

[7] I. S. Cardenas and J. Kim. Design of a semi-humanoid telepresence robot for plant disaster response and prevention. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2748–2753, 2019.

[8] G.S.L.K. Chand, S. Ojha, B. Johnston, J. Clark, and M.A. Williams. Fog robotics for efficient, fluent and robust human-robot interaction. *2018 IEEE 17th International Symposium on Network Computing and Applications (NCA)*, 2018.

[9] Y. Chen, Z. Du, and M. García-Acosta. Robot as a service in cloud computing. In *2010 Fifth IEEE International Symposium on Service Oriented System Engineering*, pages 151–158, June 2010.

[10] Confluent. Kafka REST Proxy. https://docs.confluent.io/current/kafka-rest/index.html.

[11] J. Dittrich and J-A Quiané-Ruiz. Efficient big data processing in hadoop mapreduce. *Proc. of the VLDB Endowment*, 5:2014–2015, 08 2012.

[12] R. Doriya, P. Chakraborty, and G. C. Nandi. 'robot-cloud': A framework to assist heterogeneous low cost robots. In *2012 Intl. Conf. on Communication, Information Computing Technology (ICCICT)*, pages 1–5, Oct 2012.

[13] Sani Nassif Farshad Firouzi, Krishnendu Chakrabarty. *Intelligent Internet of Things*. Springer, 2020.

[14] T. Field, J. Leibs, and J. Bowman. ROS Bag. http://wiki.ros.org/rosbag.

[15] Wu He, Guoqiang Yan, and L.D. Xu. Developing vehicular data cloud services in the iot environment. *Industrial Informatics, IEEE Transactions on*, 10:1587–1595, 05 2014.

[16] B. Inmon. *Data Lake Architecture: Designing the Data Lake and Avoiding the Garbage Dump*. Technics Publications, 2016.

[17] Intel IoT. More confidence, safety, and security in the digital world. https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/confidence-safety-security-in-digital-world-paper.pdf.

[18] Semi Koen. Architecting a Machine Learning Pipeline. https://towardsdatascience.com/architecting-a-machine-learning-pipeline-a847f094d1c7, 04 2019.

[19] A. Kristoffersson, S. Coradeschi, and A. Loutfi. A review of mobile robotic telepresence. *Adv. in Hum.-Comp. Int.*, 2013, January 2013.

[20] James KUFFNER. Cloud-enabled humanoid robots. *Humanoid Robots (Humanoids), 2010 10th IEEE-RAS International Conf. on, Nashville TN, United States, Dec.*, 2010.

[21] Advanced Telerobotics Research Lab. ATR Lab: Big Robot Data. http://www.atr.cs.kent.edu/research/big-robot-data/.

[22] Marvin Minsky. Telepresence. *OMNI*, June 1980.

[23] L. Muratore, B. Lennox, and N. Tsagarakis. Xbotcloud: A scalable cloud computing infrastructure for xbot powered robots. 10 2018.

[24] Real-Time Innovations (RTI). RTI DDS Connext. https://www.rti.com/products/connext-dds-professional.

[25] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu. Edge computing vision and challenges. *IEEE Internet of Things Journal*, Oct 2016.

[26] Spark. Spark. https://spark.apache.org/docs/2.1.0/streaming-kafka-0-8-integration.html.

[27] N. Tian, M. Matl, J. Mahler, Y. X. Zhou, S. Staszak, C. Correa, S. Zheng, Q. Li, R. Zhang, and K. Goldberg. A cloud robot system using the dexterity network and berkeley robotics and automation as a service (brass). In *2017 IEEE International Conf. on Robotics and Automation (ICRA)*, pages 1615–1622, May 2017.

[28] R. Toris, J. Kammerl, D. V. Lu, J. Lee, O. C. Jenkins, S. Osentoski, M. Wills, and S. Chernova. Robot web tools: Efficient messaging for cloud robotics. In *2015 IEEE/RSJ International Conf. on Intelligent Robots and Systems (IROS)*, pages 4530–4537, Sep. 2015.

[29] K. M. Tsui, M. Desai, H. A. Yanco, and C. Uhlik. Exploring use cases for telepresence robots. In *Proc. of the 6th Intl. Conf. on Human-robot Interaction*, HRI '11, pages 11–18, New York, NY, USA, 2011. ACM.

[30] M. Waibel, M. Beetz, J. Civera, R. D'Andrea, J. Elfring, D. Gálvez-López, K. Häussermann, R. Janssen, J. M. M. Montiel, A. Perzylo, B. Schießle, M. Tenorth, O. Zweigle, and R. V. De Molengraft. Roboearth. *IEEE Robotics Automation Magazine*, 18(2), June 2011.

[31] J. Wan, S. Tang, H. Yan, D. Li, S. Wang, and A. V. Vasilakos. Cloud robotics: Current status and open issues. *IEEE Access*, 4:2797–2807, 2016.

[32] X. Xudong and G. Rui. Research on storage and processing of mongodb for laser point cloud under distribution. In *3rd international conference on materials engineering, manufacturing technology and control*, pages 1559 – 1564. Atlantis-press, 2016.

[33] YouTube. Autonomous driving by cruise lead. https://youtu.be/s-8cYj_eh8E.

[34] M. Zaharia, M. Chowdhury, M. J. Franklin, S. Shenker, and I. Stoica. Spark: Cluster computing with working sets. In *Proceedings of the 2nd USENIX Conf. on Hot Topics in Cloud Computing*, HotCloud'10. USENIX Association, 2010.

A Vision based Voice Controlled Indoor Assistant Robot for Visually Impaired People

Saifuddin Mahmud
Advanced Telerobotics Research Lab
Computer Science
Kent State University
Kent, Ohio, USA

Redwanul Haque Sourave
Computer Science and Engineering
Khulna University of
Engineering & Technology
Khulna, Bangladesh

Md. Milon Islam
Computer Science and Engineering
Khulna University of
Engineering & Technology
Khulna, Bangladesh

Xiangxu Lin
Advanced Telerobotics Research Lab
Computer Science
Kent State University
Kent, Ohio, USA

Jong-Hoon Kim†
Advanced Telerobotics Research Lab
Computer Science
Kent State University
Kent, Ohio, USA

***Abstract*—Due to ever-increasing number of visually impaired people, the demands of assistant robots have been growing significantly. For the visually impaired, "independent living" which is the key to rebuild the dignity and self-confidence that can be achieved by assistant robot. This paper describes the design and development of a personal assistant robot, which is controlled by voice commands to find and determine the relative location of the object in the indoor environment with a co-relation factor (CRF) algorithm developed in this study. The proposed assistant robot is a semi-humanoid robot consists of several cameras on the different location of its body. The cameras are used for autonomous movement, object detection, distance measurement, and motion planning. In addition, the robot keeps the user informed about the outcomes of its actions increases the usefulness of the robot. To justify the proposed system and measure its performance, successful experiments were conducted in indoor environments. The experimental results demonstrate that the proposed assistant robot performs all of its actions with high accuracy which will make visually impaired people feel the indoor environment safe, convenient and comfortable.**

***Index Terms*—Assistive Technology, Object Detection, Visually Impaired People, Voice Control, Robot Vision, Motion Planning, YoloV3.**

I. Introduction

With the rapid increase of total population, the number of visually impaired people are also increasing significantly. According to World Health organization, there are almost 253 million individual are visually impaired all over the world where 36 million are totally blind and the rest 217 million individuals have some visual problems [1]. Vision impairment is indeed a very frequent and unendurable physical disability among several disorders. A person with visual disabilities faces numerous challenges like mobility, recognition and avoidance of obstacles, particularly in an unknown area both in indoor and outdoor [2]. Almost all cases, they have to rely on others assistance to perform their daily tasks. Although indoor environments usually have very little stringent needs, the people with visual impairments face difficulties to detect any objects owing to low vision in this scenario too.

The ongoing research is highlighting the problems associated with visually impaired people to develop assistant devices [3]. Most of the supportive devices [4] for these individuals are focused on the travel aid which is a major concern. Some of the systems like white cane and guide-dogs are commercially available for obstacle detection and avoidance both in indoor and outdoor environment. A few of the frameworks have been introduced for object detection in indoor surrounding's to assist the vision impaired individual. A deep learning based vision system developed in [5] that detected object and suggested possible path to reach the object. A smart glove proposed in [6] used for daily necessities including face recognition to identify other individual, alert signals for medications, and informing about the objects and their corresponding distances. A smartphone based object detection and classification scheme is introduced in [7] to assist visually impaired people in outdoor environment. In this system, computer vision is used for object (cars, bike, people, etc.) detection and machine learning algorithm is utilized for classification. The system proposed in [8] ensured object recognition and location estimation facility both in indoor and outdoor environment to aid visually impaired people. The developed scheme detected people, animals, vehicles in outdoor scenario, and chair, table, computer monitor in indoor environment successfully. A voice controlled assistive robotic system is demonstrated in [9] for physically disabled individual to detect and pick up particular object. A smart stick [10] is developed to assist visually impaired individual for leading normal life by providing the facility of obstacle and pothole detection, and object recognition using cloud API. The potential efforts have been made to develop supportive tools to aid visually impaired people so that they can lead their life like a normal person.

This paper aims to assist the visually impaired people by

(†)The corresponding author's email: jkim72@kent.edu

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

finding the necessary objects using a physical robot through voice command in indoor environment. Initially, a deep learning architecture (YoloV3) is implemented on the physical robot for object detection and OCR for room identification. While the robot is trained with several objects of indoor environment, voice command is given to the robot through google assistant for finding users' required object. With the voice command, the physical robot finds out the target object along with reference object and gives the necessary relative location of the object to the user successfully. In overall, the physical robot acts as a personal assistant for the people with visual impairment in indoor environment.

The remaining part of the paper is organized as follows. Section II demonstrates the most recent works that are developed for assisting visually impaired people in real-world scenario. The overall system design comprising of physical robot, object detection, and voice control as well as feedback are described in section III. Section IV depicts the experimental results of target and reference object detection using deep learning technique according to the voice commands. Lastly, Section V concludes the paper.

II. Overview of the State of the Art

Many organizations have been introduced different types of assistant devices to assist the visually impaired people in indoor environments. Various technologies like computer vision, sensory systems, robotics, smartphone technology, and artificial intelligence are used to develop supportive tools for these types of people. The works associated with this field are outlined as follows.

Lin et al. [11] developed an assistant system for visually impaired people in indoor environment utilizing deep learning architecture (Faster-RCNN). The system detected the daily necessary components like fan, chair, desk, and dehumidifier in the home environment and notified the users using smartphone with the current location of the object along with the movement location. The highest precision rate obtained by the system is 96.9 % for the dehumidifier. The computational cost for the developed system is quite low. Khairnar et al. [12] proposed a hand-held assistant tool named "PARTHA" for visually impaired users in indoor environment for object recognition. The system is able to provide the facility of obstacle detection and avoidance, object recognition, and real-time location of objects. In this system, smart glove and smartphone are integrated into a single module. The system detected the cell-phone, chair, bottle, any person in indoor environment. However, the location between the users and objects is set manually by the authority whereas the shortest path finding may be the further solution. A voice controlled individual support robot for elder people is developed in [13]. The communication between the robot and the user is done using a smartphone through Bluetooth channels. The robotic system is comprised of a camera and robotic arm where the camera is utilized for object detection and the arm is used for picking up the object. For holding any object, the gripper is used where a force sensitive resistor is integrated with the gripper to perform the task. In this framework, Yolo algorithm is utilized for object recognition. The system detected the indoor objects like bottle successfully.

Pramanik et al. [14] developed an assistant robot targeting the physically challenged people that is worked based on voice command in the hospital environment. The developed tool is comprised of several components like Arduino Mega, Raspberry Pi, robotic hand, Veear module, and VR command module. The robotic system grabbed and picked up the object according to the voice command. The full cost for developing the system is approximately $360. Patel et al. [15] introduced an object detection system for the people having visual impairment in indoor environment which is worked based on multi-sensor data. Several electronic components like USB camera, IR sensor, ultrasonic sensor, and Raspberry Pi are attached with a stick. The developed prototype detected various indoor objects like chair, table, and so on. The user is notified with the presence of any objects through a buzzer module. Reddy et al. [16] proposed a system operated by voice command for the physically challenged individual that is able to detect obstacle, smoke and fire in the case of emergency. The command from the users is sent through "voicebot" app. The developed system used Arduino UNO, ultrasonic sensor, Bluetooth module, smoke detector, fire detector, and Wi-Fi module for prototype development. The prototype is developed highlighting the principle of wheelchair but the full development is not completed. However, the used sensors worked in a certain ranges, cannot cover a wide area. Nguyen et al. [17] introduced a vision based framework to aid the visually impaired people in indoor environment. The proposed system gave a possible map from user to a particular object in building, school, or hospital environment. The users are suggested to follow an assistant robot to navigate for reaching the object. However, the system is only applicable for small pervasive surroundings but cannot perform in loop environment.

Vlaminck et al. [18] developed an indoor assistant scheme for the vision impaired individual that is highlighted to localization, and obstacle detection and classification. The localization is confirmed with depth and color information which is captured through a RGB-D camera. The developed system categorized the objects as walls, steps, doors, etc. The computational cost for the proposed scheme is comparatively low than the existing systems. The authors of [19] proposed a system that is able to detect objects along with moving guidelines in real-time highlighting the visually impaired. In this system, Microsoft Kinect is used a capturing device. For the experiment, a total of 600 samples are considered in the developed system. The framework detected the moving and still objects from video with the accuracy of 92%, and 87% respectively. However, the developed system cannot well perform in direct sunlight environment due to the use of Kinect sensor. Irvin et. al. proposed a telepresence robot [20] which can be used as a personal assistant robot for the disable people. This robot has it's own autonomy and it can be controlled by the remote operator. Tepelea et al. [21] presented a vision module for the visually impaired individual

that ensured the facility of object detection both in indoor and outdoor environment. Template matching is used as a feasible technique in Raspberry Pi platform. The developed system provides the results with low computational time successfully. However, the illumination effect is not considered here which is very important for real-life object detection.

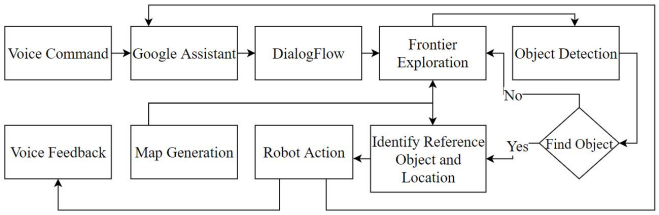


Fig. 1. Overall system flow of the developed assistant robotic system

III. System Design

The proposed assistant robot is mainly aiming at efficient and reliable approach towards designing and deploying an assistant system that includes object finding and detection of the relative location of the object functionalities to aid a visually impaired person in indoor environment. An overview of the system is depicted in Figure 1. For voice command and feedback, google assistant is used while Robot Operating System (ROS) is responsible for the coordination among all the modules. The proposed system is mainly consists of five different modules:

- Psychical Robot
- Voice Control
- Path Planning
- Object Finding
- Voice Feedback

A. Psychical Robot

In our proposed system, TeleBot-R2 [22] is used as a physical robot. This is a semi-humanoid robot was designed and developed at our lab [23]. It has a eye-catching design Figure 2 and it is equipped with 360 degrees of maneuverability with a multi-track functionality. It has a tank base with additional flippers which helps the robot to move on staircases if there are multiple floors. For tracking and depth information for path planning, one Astra S camera is attached on the forehead position. A 4k RGB camera is attached to the head of the robot for object detection. The robot is able to rotate the RGB camera to capture wide area.

B. Voice Control

As the visually impaired people do not have the ability to see around them, it is difficult for them to operate the robot by using any other devices. By using voice command, this limitation can be overcome.

Google assistant is used to take the user voice input and process them in order to command robot. To make the process easier, Dialogflow, a natural language understanding platform provided by google for interactive voice response is used. A

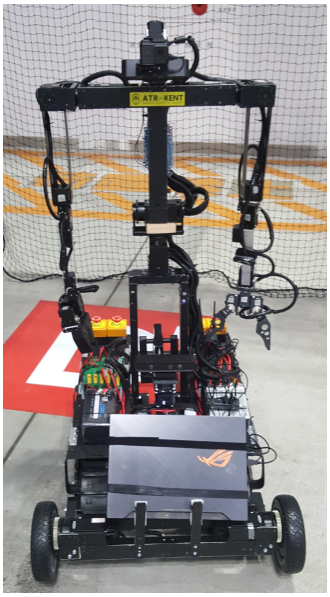


Fig. 2. The Physical Robot: Telebot-2

set of predefined command, such as ”find me my phone”, ”where is my cup” and etc. are set in our system. By using the action function and define the entity type in DialogFlow, our system will automatically separate the target object from trigger sentence. The user will not be constrained by what sentence to say as the AI based Dialogflow maps with the corresponding built dialogues to provide exact command to the robot. In our proposed system, open source google assistant API is used in this purpose.

C. Path Planning

Path planning is one of the most important module of our system as autonomous movement of the robot and overall command execution time depend on it. We have used “frontier-exploration” package of ROS which is a frontier-based approach for autonomous exploration [24]. Using the depth camera, robot collect new information about the indoor world and develop a map autonomously. Our path planning module allows the robot to scan the environment thoroughly, build a map, identify the obstacles and frontiers in the map, and continue explore those unknown areas until all the areas are fully explored and mapped. Once the map of the indoor environment is fully built, the robot performed very efficiently. During the execution of a command when robot find a specific object then it store the location of the object for future use.

D. Object Finding

Once the robot gets the command to find a specific object through Google Assistant then it navigates in the indoor environment to find the object. As the visually impaired people has limited vision, they need some relative information about the position of the object. To provide the proper location information of the object, the robot needs to perform the following actions.

1) Target Object Detection
2) Reference Object Detection
3) Room Identification

For the target and reference object detection , we have used YoloV3 [25], a convolutional neural network(CNN). We have trained YoloV3 using our own dataset on target object(ex: mask, sunglasses), and reference object(table, sofa). For room identification we have read-text package which is based on OCR. To prepare the datasets for the target object and reference object, we have collected 1500 images per object.

Then, we have manually labeled the image by drawing a bounding box around the object and stored them using the following format <class> <x-center> <y-center> <width> <height>. All the values are normalized in the range of 0 to 1. For training the YoloV3 network, we have used Google COLAB [26], a Machine Learning Research Tool, which is a free cloud service and supports GPU for free. After the training process is done, we have deployed it to the robot for for obtaining the relative location of an object by the result of target object detection, room identification and reference object detection. When objects are detected YoloV3, provides the bounding boxes around them, name of the object and level of confidence when the object was detected. The area of the boxes, distance between the bounding boxes and level of confidence are used to determine the co-relation among them. To determine the co-relation and provide relative location information, we have considered the following cases:

- Case 1: The target object is detected, one reference object is detected and room is identified. The bounding box of the target object is on the right/left side of the reference object's bounding box. In Figure 3 a) Target object, (water-bottle) is at the right side of the reference object(chair).
- Case 2: The target object is detected and multiple reference objects are detected and room is also identified. The box of reference object which has the minimum distance form the target object box will be used for providing the reference. In Figure 3 b) the target object, mobile phone, is detected and sofa and chair are detected as reference objects. As chair is more closer to the mobile phone than sofa, robot will use chair a reference object.
- Case 3: Room is identified and the bounding box of the target object is inside the bounding box of the reference object. Figure 3 c) shows that target object (mask) is on the reference object(sofa)
- Case 4: Room is identified, object is detected but there is no reference object. In figure 3 d) there is no reference object near the target object(key)

E. Voice Feedback

When the object finding process is done, the robot send the relative location of the object through google assistant by voice feedback. As the "Dialogflow" is used, the robot can send the voice feedback in a natural way. For, example according to the Figure 3 a) after detection of the target and reference object, and identification of room robot will provide the relative location the object and the Google Assistant will generate "Water-bottle found. It is on the right side of the chair in the bedroom"

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

To evaluate the performance of the proposed system, experiments were conducted inside a one bedroom apartment. We placed several target objects at bedroom and living room with different reference objects similar as Figure 3. The Table 1 shows that sunglassess, water-bottle, mask, phone and key ring are taken as target object whereas bed-table, chair and sofa are taken as reference object. The robot is able to detect the target and reference object successfully to provide the relative location by using its vision and co-relation factor described in Figure 3. The overall performance of the proposed system is satisfactory but in the bedroom performance degrades.

Due to the lack of illumination , illumination inequality our system drops its performance. However, The results are very promising even though experiments with the physical robot in real-world scenario were limited due to the current COVID-19 pandemic and those challenges can be overcome by combining the depth information of an object during object detection which is one of our future research plan. Due to the limitation of YOLO algorithm which was based on 2D images, the confidence level during object detection varied due to the change of the distance which is show in Figure 4.

V. CONCLUSION

The proposed system presented a vision based voice controlled personal assistant robot to aid visually impaired people in indoor environment. In this system, the physical robot detects the target object along with reference object successfully according to the voice command. The assistant robot is trained with several indoor objects using deep learning network (YoloV3) for object detection and OCR for room identification. Google assistant service is utilized for voice input (for finding particular object) from user to robot and voice feedback (for providing object location with reference object) from robot to user. It is evident from the experimental findings that the proposed system effectively detects mask, sunglasses, water-bottle, phone, etc. in indoor space. The experimental outcomes reveal that the proposed scheme is effective, reliable and fulfills all the targeted aims to help people with visual disability. The developed system only provides the relative location of the target object ; picking up the object from the current position and provide it to the user would be the effective solution as a future work.

REFERENCES

[1] R. R. Bourne, S. R. Flaxman, T. Braithwaite, M. V. Cicinelli, A. Das, J. B. Jonas, J. Keeffe, J. H. Kempen, J. Leasher, H. Limburg *et al.*, "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis," *The Lancet Global Health*, vol. 5, no. 9, pp. e888–e897, 2017.

[2] M. M. Islam, M. S. Sadi, and T. Bräunl, "Automated walking guide to enhance the mobility of visually impaired people," *IEEE Transactions on Medical Robotics and Bionics*, pp. 1–1, 2020.

Fig. 3. a) Target object is at the right side of the Reference object b)Multiple reference object detected c)Target object is on the Reference object d) Target object detected without Reference

TABLE I
OVERALL PERFORMANCE OF THE PROPOSED SYSTEM

Target Object	No. of trial	Reference Object at Bedroom	Success rate of Target object Detection in Bedroom	Success rate of Reference object Detection in bed Room	Reference Object at Bedroom	Success rate of Target object Detection in Living	Success rate of Reference object Detection in Living Room	Average Accuracy(%) For Relative location information Extraction
Sunglasses	20	Bed-table	19	18	Sofa	20	19	90
Water-bottle	20	chair	20	19	Chair	20	19	95
Mask	20	Bed-table	19	18	Sofa	20	19	90
Phone	20	Chair	19	19	Chair	20	19	95
Key Ring	20	X	19	X	X	20	X	95

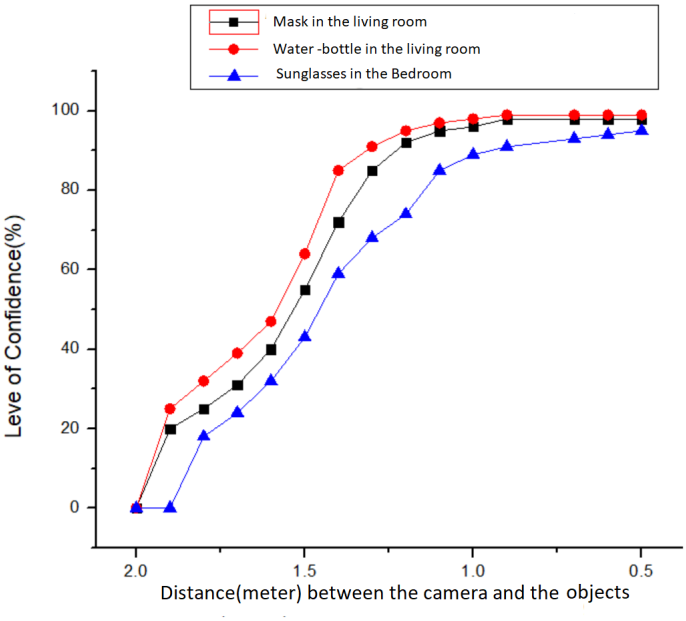


Fig. 4. Distance vs Level of Confidence in object detection

[3] M. M. Islam, M. Sheikh Sadi, K. Z. Zamli, and M. M. Ahmed, "Developing walking assistants for visually impaired people: A review," *IEEE Sensors Journal*, vol. 19, no. 8, pp. 2814–2828, 2019.

[4] R. Tapu, B. Mocanu, and T. Zaharia, "Wearable assistive devices for visually impaired: A state of the art survey," *Pattern Recognition Letters*, 2018.

[5] M. Shih, Y. Chen, C. Tung, C. Sun, C. Cheng, L. Chan, S. Varadarajan, and M. Sun, "Dlwv2: A deep learning-based wearable vision-system with vibrotactile-feedback for visually impaired people to reach objects," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1–9.

[6] L. González-Delgado, L. Serpa-Andrade, K. Calle-Urgiléz, A. Guzhñay-Lucero, V. Robles-Bykbaev, and M. Mena-Salcedo, "A low-cost wearable support system for visually disabled people," in *2016 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, 2016, pp. 1–5.

[7] R. Tapu, B. Mocanu, A. Bursuc, and T. Zaharia, "A smartphone-based obstacle detection and classification system for assisting visually impaired people," in *2013 IEEE International Conference on Computer Vision Workshops*, 2013, pp. 444–451.

[8] P. Rohit, M. S. Vinay Prasad, S. J. Ranganatha Gowda, D. R. Krishna Raju, and I. Quadri, "Image recognition based smart aid for visually challenged people," in *2019 International Conference on Communication and Electronics Systems (ICCES)*, 2019, pp. 1058–1063.

[9] R. A. John, S. Varghese, S. T. Shaji, and K. M. Sagayam, "Assistive device for physically challenged persons using voice controlled intelligent robotic arm," in *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2020, pp. 806–810.

[10] D. Bharatia, P. Ambawane, and P. Rane, "Smart electronic stick for visually impaired using android application and google's cloud vision," in *2019 Global Conference for Advancement in Technology (GCAT)*, 2019, pp. 1–6.

[11] W. Lin, M. Su, W. Cheng, and W. Cheng, "An assist system for visually impaired at indoor residential environment using faster-rcnn," in *2019 8th International Congress on Advanced Applied Informatics (IIAI-AAI)*, 2019, pp. 1071–1072.

[12] D. P. Khairnar, R. B. Karad, A. Kapse, G. Kale, and P. Jadhav, "Partha: A visually impaired assistance system," in *2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA)*, 2020, pp. 32–37.

[13] J. U. K, I. V, K. J. Ananthakrishnan, K. Amith, P. S. Reddy, and P. S, "Voice controlled personal assistant robot for elderly people," in *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, 2020, pp. 269–274.

[14] S. K. Pramanik, Z. A. Onik, N. Anam, M. M. Ullah, A. Saiful, and S. Sultana, "A voice controlled robot for continuous patient assistance," in *2016 International Conference on Medical Engineering, Health Informatics and Technology (MediTec)*, 2016, pp. 1–4.

[15] C. T. Patel, V. J. Mistry, L. S. Desai, and Y. K. Meghrajani, "Multisensor

- based object detection in indoor environment for visually impaired people," in *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2018, pp. 1–4.

[16] P. M. Reddy, S. P. Kalyan Reddy, G. R. Sai Karthik, and B. K. Priya, "Intuitive voice controlled robot for obstacle, smoke and fire detection for physically challenged people," in *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*, 2020, pp. 763–767.

[17] Q. Nguyen, H. Vu, T. Tran, and Q. Nguyen, "A vision-based system supports mapping services for visually impaired people in indoor environments," in *2014 13th International Conference on Control Automation Robotics Vision (ICARCV)*, 2014, pp. 1518–1523.

[18] M. Vlaminck, L. H. Quang, H. Van Nam, H. Vu, P. Veelaert, and W. Philips, "Indoor assistance for visually impaired people using a rgb-d camera," in *2016 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, 2016, pp. 161–164.

[19] A. N. Zereen and S. Corraya, "Detecting real time object along with the moving direction for visually impaired people," in *2016 2nd International Conference on Electrical, Computer Telecommunication Engineering (ICECTE)*, 2016, pp. 1–4.

[20] D. Y. Kim, I. S. Cardenas, and J.-H. Kim, "Engage/disengage: Control triggers for immersive telepresence robots," in *Proceedings of the 5th International Conference on Human Agent Interaction*, ser. HAI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 495–499. [Online]. Available: https://doi.org/10.1145/3125739.3132603

[21] L. Tepelea, I. Buciu, C. Grava, I. Gavrilut, and A. Gacsádi, "A vision module for visually impaired people by using raspberry pi platform," in *2019 15th International Conference on Engineering of Modern Electric Systems (EMES)*, 2019, pp. 209–212.

[22] I. S. Cardenas and J. Kim, "Design of a semi-humanoid telepresence robot for plant disaster response and prevention," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 2748–2753.

[23] [Online]. Available: http://www.atr.cs.kent.edu/

[24] B. Yamauchi, "A frontier-based approach for autonomous exploration," in *Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97. 'Towards New Computational Principles for Robotics and Automation'*, 1997, pp. 146–151.

[25] J. Redmon and A. Farhad, "Yolov3: An incremental improvement, arxiv, 2018."

[26] https://colab.research.google.com/notebooks/intro.ipynb.

Smart temperature-controlled infant car seat using thermoelectric devices

Roshni Senthilkumar
Electrical and Electronic Engineering
Heriot-Watt University.
Dubai, United Arab Emirates
rk49@hw.ac.uk

Mohamed Al Musleh
Automotive Engineering
Heriot-Watt University.
Dubai, United Arab Emirates
m.al-musleh@hw.ac.uk

***Abstract*—The thermal comfort for children and adults inside the car is a crucial issue to improve as the hot environmental conditions can be irritable and uncomfortable. Being inside an extremely hot car for a prolonged time can lead to major health problems. This paper focuses on the thermal comfort of infants. Traditionally air conditioning systems are used to cool down an infant car seat which is not effective due to the portability and cooling duration. In this paper, heat relief for the child is researched using a thermoelectric cooling system. This paper reviews various factors concerning the development of a temperature-controlled infant car seat. This study focuses on the health and safety of children inside the cars, the subjective needs of the parent, heat management in the car cabin environment, thermoelectric modeling, and the methods for precise temperature control for effective thermoelectric cooling. In-depth research has been conducted on the contributions of other researchers on the thermoelectric cooling system in a car. The thermoelectric device in the infant car seat was tested in thermal simulation.**

Keywords—thermoelectric device, PID control, monitoring system, thermal comfort, smart control

I. INTRODUCTION

Child safety and comfortability is an important factor to be considered in an infant car seat. In the last few decades, thermoelectric cooling (TEC) has gained significant popularity in the automotive industry for car seat coolers because of their efficient, compact, and noiseless operation. [1] Additional advantages include stationary operation, no operating fluids nor chemical reactions, and do not produce emissions. This makes TECs advantageous compared to conventional ventilation systems. Although ventilation systems for car seats have been used widely in the automotive industry such as in Tesla [2], Ford [3], Marlow automotive [4], this project focuses to develop a ventilation system for the detachable child car seats which is not frequently found in the market.

A. Market Demand

In recent years, the demand for the baby car seat is increasing all over the world. According to the research by *Technavio*, a leading global technology research center, there is a huge need for infant car seats especially, the infant car seat for 6 to 24 months old child [5]. As many people spend money on purchasing cars, more people can be seen traveling on the road. So, the car seats must be more comfortable in terms of instant cooling and heating features and the ability to monitor their child. A survey was conducted by the author to show the need for a cooling system in an infant car seat by parents. The demo-graph for the survey were parents living in the Middle East with around 80 survey results. The pie-chart results are shown in Fig. 1. It is seen that 83% of the survey takers felt their child was uncomfortable in the car when parked under the sun for too long. When they were asked the reason for the uneasiness, the majority of them replied that their child would sweat due to the heat inside the car. This is a significant survey result as the parents know their child best. This shows that the development of a solution to create temperature controlled infant car seats is vital in keeping children comfortable.

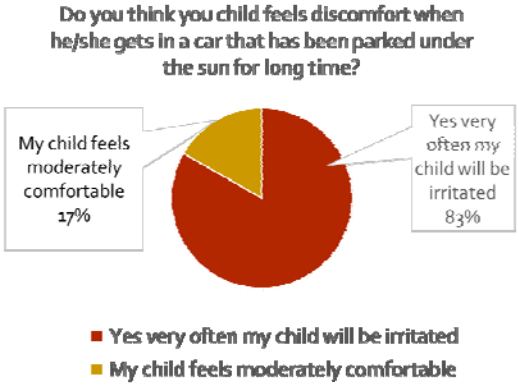


Fig. 1: Survey results.

B. Health and Safety of Children

Some standards that define the level of comfort inside the car for a human being are ISO EN 7730, ASHRAE 55-92), a CEN report (CR 1752) which shows that the temperature inside the car is extremely high which can lead to a high level of discomfort [6].

According to the journal [7], a list of body temperature benchmarks is shown in table 3.1. It is seen that the ideal body temperature of 6 months old baby is 36.5□C. According to NBCI [7], the infant must be placed at an ambient temperature of 25□C to 27□C to maintain the ideal body temperature for comfort.

The medical issues mentioned by the Health Information Research Center, people sitting in a hot car for too long can experience excessive sweating, heat strokes, hyperthermia, and HMFD (hand, mouth, foot disease). [8] However, the number of children who die from vehicle-related hyperthermia is only a small percentage, so this hazard is easily preventable. A pediatrician at UCSF children's healthcare (University of California, San Francisco) has reported that babies have a hard time regulating their body temperature, so a child's body temperature can rise 2 times faster than that of adults placing the child in a dangerous situation [9]. The heat-regulating mechanisms of the body, especially the hypothalamus regulates and maintains the body temperature according to the outside temperature. This process is called thermoregulation, where the sweating glands are activated to maintain body temperature at 37□C. In an environment of 24□C the heat loss from the body is around 146W whereas if the environment is 45□C, the heat

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

loss is -36W where the body is reducing the body heat to maintain cool body temperature [10]. For children, the rate of thermoregulation will be 2 times slower which leads to dry skin and dehydration where the body cannot absorb excessive heat. So, it is recommended that the caregiver cools the car seat before a young child is placed in it as it reduces the chance of hyperthermia and dehydration.

II. Literature Review

In this section, several elements of the cooling infant car seat are researched by examining the other authors' work.

A. Thermoelectric Cooling (TEC)

There are various thermoelectric configurations available in the market such as the High-temperature Peltier, multistage Peltier, and series-parallel Peltier [11]. High-temperature Peltier's have the integral structure to operate at high temperatures such as 200□C to provide high cooling capacity. According to this conference paper [12] where it reviews the coefficient of performance of the thermoelectric system. His results showed that the cooling mode of the system produces 33.3 W at a coefficient of performance (COP) of 0.4, and about 94.6 W of heating power at a COP of 1.15. Therefore, by using two TEC in parallel, higher COP can be achieved. TEC systems are a reliable approach that provides accurate temperature and easy installation.

B. Heat Flux inside the Car cabin

The heat flux inside the car cabin near the vicinity of the child car seat is determined by considering the thermal behavior of the passenger, vehicle cabin temperature, and the thermoelectric heat flux. According to the journal [13] the human thermal model was assumed as lumped capacitance as it can simplify complex heat flux of the body.

The amount of heat released by the individual is considered to determine the cooling load of the thermoelectric device. To find the heat load of the human on the car seat, heating pads were placed on the seat. Energy exchange between a driver's body and the automobile seat was seen [6]. Initially cabin temperature is considered at 0□C gradually the temperature was increased. While he experimented, 80W heat was generated in the heating pads from the male adult. The heat was transferred to the contacting areas and some heat dissipated to the ambient or got accumulated in the seat. However, for a child, the body contact area is smaller. Overall, it is noted that the heat flux surrounding the infant car seat is larger than the heat load of the thermoelectric. Therefore, proper heat management is introduced.

C. Heat Rejection Management

There are three types of cooling systems: conduction, forced convection, and liquid cooling. From 2016, luxury cars such as Tesla [2], Ford [3] Marlow automotive [4] have integrated liquid cooling car seats that operate using a thermoelectric cooling system. However, liquid cooling is heavy and requires complex structures which not portable for infant car seat usage.

According to the journal [14], the Peltier device was tested against the three types of the cooling system and found the heat pipe had the highest heat rejection closely followed by the heatsink system. He tested the heat sink with the integrated fan model and found that it can withstand 10W of heat flux from the Peltier device [14]. The thermal resistance of quality forced convection systems typically falls within a range of 0.02 to 0.5°C/watt [15]. The heat sink adds little weight and less stress to the thermoelectric device. Therefore, it is opted for.

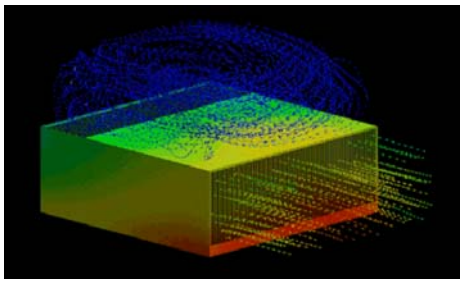


Fig. 2: CFD simulation shows a circular motion of air through a rectangular heatsink. [16]

The centrifugal fan is popularly used in the industry due to its simple design and high versatility. The airflow movement in Fig. 2 shows a high output speed. These fans can be modified to limit the income of dust as well.

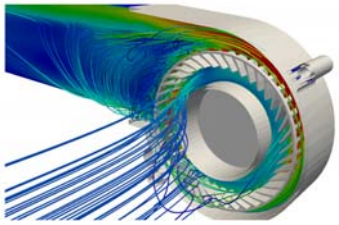


Fig. 3: Velocity streamlines for the centrifugal fan in CFD [17]

In this journal [13] [18] a heat exchanger system was designed to provide localized convective and evaporative cooling. The pipe exchanger was experimented and found that the average temperature of the air after passing through the tube exchanger was lower by an average of 2 °C. Whereas enclosure containing the TEC with twisted pipes connected to the cold and hot side of the TEC where the cold side pipe is connected to the car seat is found to be effective in keeping the airflow rate stable [19]. The air from the main cooler box will be pass through the hole to the surface allowing even distribution of air shown in Fig. 3. However, some limitations faced by the author were the high heat rejection imposed by the hot side of the Peltier device.

Moreover, the airflow rate of the fan must be studied to analyze the cooling air distribution at the surface of the seat. In the simulation conducted by Allison Bailes [20] shows that the laminar air from a wider duct area to a reduced duct area shows that the velocity at the small opening is larger. The flow velocity increased from 1m/s to 2m/s. Even after introducing the turbulent disturbance, there is a great increase in velocity. This effect is due to the Bernoulli's equation of fluid.

Moreover, it is important to keep the child comfortable. To improve the cooling effect the selection of seat cover fabric is crucial to provide comfort and cooling effect. The fabric must have high thermal conductivity and be able to absorb heat. Mesh fabric with many minuscule holes is used in many car seats to improve thermal conductivity and comfort while resting on the top. The fabrics that have high thermal resistance are cotton and cotton/polyester where treated cotton blend polyester has a thermal resistance of

0.01383 (m^2K/W) [21]. Therefore, this mesh is used in the infant car seat to improve the cooling effect. Overall, while manufacturing the infant car seat, it is suggested to use centrifugal fans and cotton fabric.

III. Methodology

The methodology followed to construct a control of the temperature of the thermoelectric device is presented. First, the heat rejection from the thermoelectric device is managed by modeling a special heat exchange unit. Then, the temperature control system is created and integrated into the microcontroller through MATLAB. The monitoring system is integrated on the seat and configured with the microcontroller. Finally, it is tested.

IV. Prototype Development

A. Thermoelectric Cooling Unit (TEC)

The implementation of the framework for heat exchange is discussed. The TEC unit is placed at the back of the infant car seat and an airflow duct is connected to the back of the seat. The airflow from the fan will pass through the air duct then to the surface of the seat. This method was selected to be simulated because the TEC unit is enclosed and safely placed at the back of the car seat.

To improve heat rejection, an active heat sink is used. A rectangular heatsink with a fan attached on top directs the hot air outside. This heatsink has a total thermal resistance of 0.179 °C/W [16]. The volumetric thermal efficiency (including fan) equals 0.021 W/°C/cc [16]. The weight of this heatsink with a fan is 194 grams which are lighter than heat pipes. The heat sink placed between two large heat sinks with an inbuilt fan. As the Peltier is placed between the heatsinks, there will be less heat stress causing less damage to the Peltier. It consists of a fan attached to a heat sink that diverts heat. The centrifugal fan used in this prototype produces an airflow rate of 14.1 CFM which is a velocity of 0.08m/s [15]. However, this velocity will be reduced as it passes through the heatsinks towards the child. A simulation was done to view the airflow that will reach the child.

The thermoelectric device TEC1-12406 is between two heat sinks. The top of the heat sink must be kept flat. The insulation materials around the TEC panel are used to reduce the escape of cold air to the surrounding environment. Insulation materials are poor thermal conductors and prevent the heat from leakage. The top heatsink collects the hot side heat and releases it to the ambient through the fan and the heat sink at the bottom. The centrifugal fan circulates the air around the heat sink to draw the heat away from the unit.

A 3D printed case is designed on the TinkerCad software. The model is created by designing a duct for the hot and cold side to pass into, as seen in Fig. 5 and 6. This layout has a space for both the hot side convective air and the cool side convective air to move through. The air movement from the hot side passes through the gap and cools the heatsink. The circular gap is given for four cables of the thermoelectric device.

The working of the heat rejection model is explained. The thermoelectric devices were placed inside a 3D printed enclosure and kept on the backside of the infant car seat. The thermoelectric device stays between two heat sinks with a forced convection fan. The insulation materials around the TEC panel are used to reduce the escape of cold air to the surrounding environment. Insulation materials are poor thermal conductors and prevent the heat from leakage. The cool air is sent to the child through an airflow pipe which is connected between the cooling unit and the infant car seat.

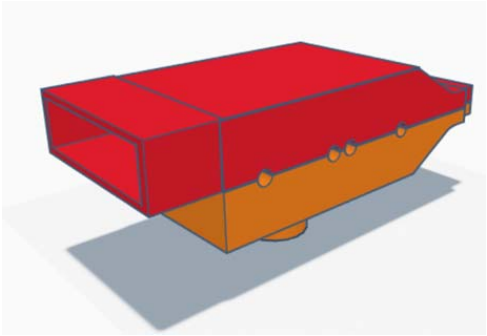


Fig. 5: 3D model of the thermoelectric unit

Fig. 6: 3D printed case

B. Temperature Controller Modelling

The temperature control system of TEC is achieved using PID control. The mathematical model is assumed in ideal conditions that there is no leakage heat flux and the surrounding temperature is constant. The hot side of the TEC is kept at a constant temperature using a heatsink. Assuming the cold side does not absorb any heat Qc=0. The cold side temperature in terms of current I am derived by modifying equation 1 to get Tc. The non-linear characteristic of the thermoelectric device is seen. This equation is used to model the thermoelectric device. [8] [16]

$$T_c = \frac{0.5I^2R_e + K_{th}T_h + Q_c}{SI + K_{th}} \quad (1)$$

This aspect can be validated by graphing the temperature of the cold side. The parameters for the Peltier device TEC1-1206 from the datasheet are T_h =323K, R_e =2.3Ω. The thermal conductance and Seebeck coefficient K_{th}=0.537, S=0.0514. It is seen that the cold side temperature decreases until the equilibrium point of 5.944A and -7.004□C. When the current is increased further beyond the equilibrium point, the cold temperature rises which is not an ideal operation and resembles instability. The Peltier releases heat instead of absorbing the heat. The relationship between current and temperature is non-linear.

The non-linear characteristic can be dealt with by developing a control system that limits the current supplied based on the performance curve of TEC.

The H-bridge is used to control the thermoelectric unit. The H-bridge provides a bidirectional current drive to the

TEC. The magnitude of the current is controlled by regulating the duty cycle and frequency of the pulses produced by MOSET. It was built in Pspice and simulated to find the output voltage and current at the TEC. The input voltage was set at 10V with a duty cycle of 0.82 with a period of 6.66ms.

$$V = D \times V_{in} = 0.82 \times 10 = 8.2V \quad (2)$$

The voltage at the TEC must be maintained at 8.2V. The PWM signal was placed only on switch 1 and switch 4, therefore, the forward current will only pass through the TEC.

C. *Smart Monitoring System for Infant Car Seat*

The camera was mounted on the handle of the infant car seat to achieve a perfect view of the child. This was achieved by measuring the size of the camera and drilling a hole on the handle as shown in Fig. 7. The wires were fed through the handle and connected to the electronics. The display of the camera is shown in Fig. 8. This display will be placed on the dashboard of the car. When the caregiver is driving, and the baby seat is kept at the rear seat, the caregiver can view the child and its movement through this display. This is a wired camera with a wired display which is connected to the battery power source. However, the long wires of the display made it hard to mount the display on the dashboards. This can be overcome by using a Bluetooth camera and display.

Fig. 7: Infant Camera

Fig. 8: Integrate monitoring display installed in the seat.

The remote start allows the user to switch on the cooling system before entering the car. The infant car seat will be cooled and ready to use. This was achieved through transmitters and receivers connected to the power supply. It uses RF (radiofrequency) 2.5GHz wireless signals to communicate between the receiver and transmitter. In the Fig. 9 below, the remote has two buttons, pressing the A button, turns on the power supply powering the whole system while pressing button B, the whole controller turns off. The wireless range of the transmitter is 1.5m.

Fig. 9: Remote Controller

D. *Electronic Control Unit*

The electronic unit can be divided into 4 parts: Power supply Unit, temperature control unit, detection unit, and user interface unit. The overview of the electronic design is shown in Fig. 10.

The electronic circuit diagram is centered around the Arduino ecosystem due to its accessible and easy user interface. The power supply unit provides power to all the electronics by regulating the DC input voltage of 12V from the car battery. The RF remote start is connected to the power supply. The monitoring camera system and the microcontroller is connected to the remote start. The H-bridge is connected to the Peltier which will ensure there is enough power to drive the Peltier device. The temperature control unit uses a PID feedback system to automatically control the Peltier device within the required temperature using the Arduino microcontroller. The user interface contains the monitoring system with a camera and the Bluetooth connect app. The detection unit senses the ambient and temperature of the TEC.

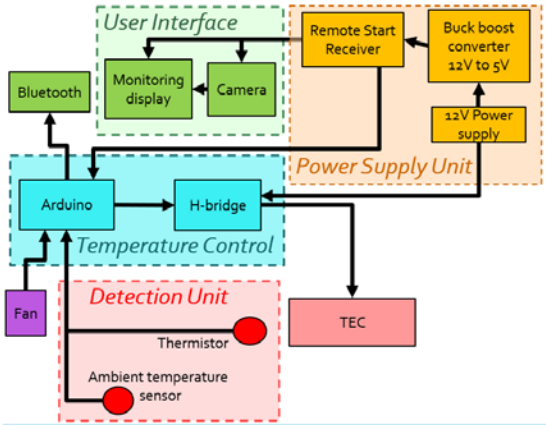


Fig. 10: Block diagram of the electrical connections.

The digital design of the temperature control system is summarized as follows. It is done automatically by the microcontroller using the PID control system. First, the temperature of the ambient temperature neat the child, the cold side temperature, and the hot side temperature is sensed, every 10s. When the system turns on, the temperature is sensed and transmitted to the microcontroller. If the temperature exceeds the set temperature, the cooling system turns on and if the temperature goes lower than the set temperature, the heating function turns on. The temperature can be either set by the user or the user can toggle between the modes. The mode can be toggled using the user interface on the phone. This code is given in the appendix A. The sensed cold side temperature is converted into a corresponding current value using the TEC characterization

module. The corresponding PWM of the current is obtained and sent to the TEC. If the ambient temperature is reaching above 35□C, an alert is sent to the phone. If the cold side temperature is increasing due to malfunction, then the power input the TEC is stopped unit the cold side temperature reduces to ambient temperature. This function will be viewed using the TEC characterization equation.

The temperature is controlled with a set power to the operator at stable TEC conditions. The power-limited does not allow temperature more than the rated power of the TEC. Moreover, the TEC can be switched off by turning off the enable switch. By doing this the power will not be provided to the TEC.

The Bluetooth module to the Arduino and the protocol step was achieved, it can be displayed on the phone. An online software called RemoteXY along with its supporting app on the phone is used to display the temperature reading. The Bluetooth connection is activated on the phone and paired with the app. The app layout is given in the following Fig. 11. It displays 3 features; temperature monitoring, temperature control setting, and alert system. When the TEC system is switched on, the live temperature readings are displayed. If the temperature reached above the threshold, the alert system is activated. This app is essentially created for the caregiver of the child to use, to adjust according to the child's comfort. The alerts and user set temperature features to allow for easy monitoring of temperature. The app was tested, and the temperatures were changed upon changing the cooling levels. When a high temperature was detected, the alert was blinking.

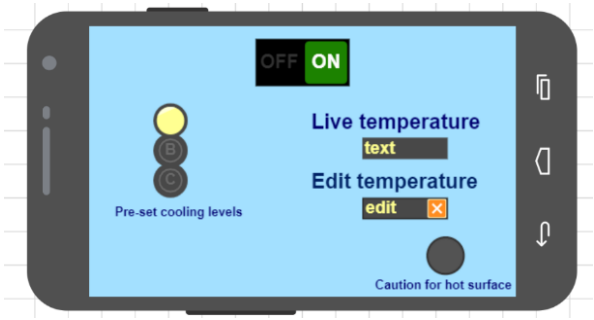


Fig. 11: Developed the user interface on the phone

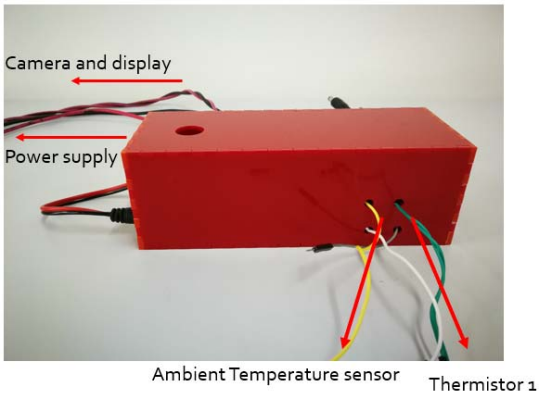


Fig. 12: Enclosure for the electronics

The electronics of the control system are kept inside the enclosure as shown in Fig. 12. The output wires to the camera, display, and power supply are seen. The temperature sensor for TEC and ambient temperature is also shown.

E. Integration the complete infant car seat

The cooling module consists of the thermoelectric device, heat sinks, 24V fan enclosed inside a 3D printed case. An air duct is fed through the infant car seat to which the cooling unit is mounted. Especially sewn fabric at the backrest improves heat conduction. This is shown in Fig. 13. A video link done by the author shows the built and attached components of the infant car seat. [22]

Fig. 13: Completely integrated seat with the special fabric.

V. Results and Discussion

The vehicle cabin temperature was found by conducting temperature tests inside the car. Four tests were conducted: parking the car under the sun with windows closed and windows opened and the car under the shade with windows closed and opened. The temperature logger Tiny Tag was placed on the back seat of Honda and Jeep and tested. The control variable is the temperature and content variables are the time of the day. The test was conducted for 24 hours but the data is compared for the time from 11 am till 1 pm.

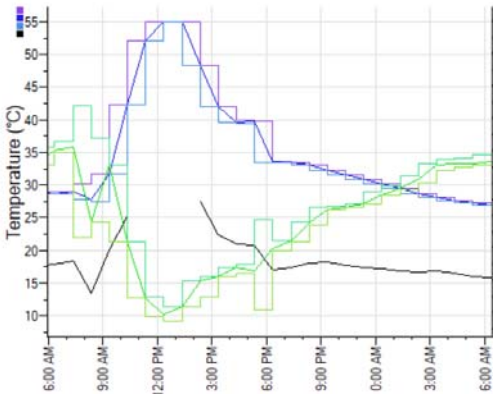


Fig. 14: Temperature tests with windows closed (Blue line: Temperature; light green: Humidity; Black: Dew point.)

It is found from Fig. 14 that the temperature inside a car parked under the sun reached considerably high temperatures, such as about 55 □C at the rear seat in 90 mins, when the outside temperature is around 32□C.

Modeling the heat distribution on the surface of the car seat requires three elements which are TEC, infant, and insulation. This is simulated in energy2D to demonstrate the heat distribution around the infant car seat. The TEC module is defined by a hot block of 40 °C and a cold block of 20 °C reflecting a heat source. The TEC is sandwiched between two heat sinks with thermal conductivity of 167 W/mK. The ambient temperature of the car was set at 45□C. The fan next to the heatsink has a wind speed of 2 m/s. The child is represented by a 36.5 °C heat source. The car cabin is concealed inside a boundary. Fig. 15 displays the effects of the simulation which shows the movement of the heat flux according to this model. A video link that shows the movement is attached in the reference. [23]

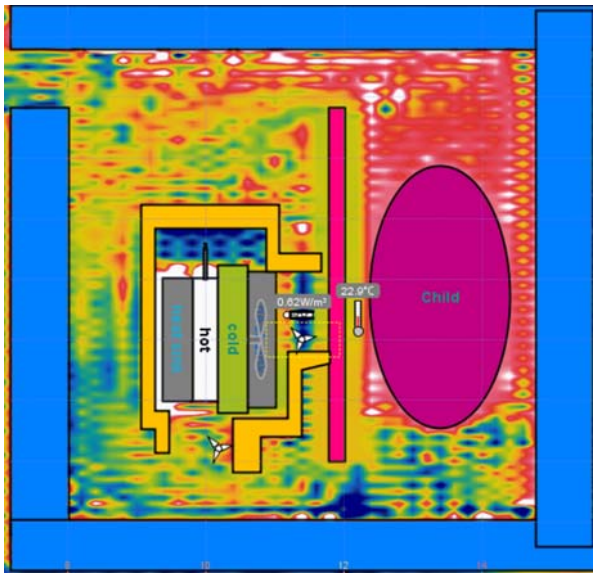


Fig. 15: Simulation result of the integrated cooling system. [23]

The simulation reveals that the temperature around the occupant is lowered to 21□C, which is the temperature set on the cold side of the TEC. The red and white dots depict hot temperatures, while the black and the blue dots depict cold temperatures. It can be seen that cold air moves towards the seat cover material, and then towards the infant.

To assess the validity of the design results, an experiment was conducted by uploading the code into Arduino. The integrated cooling unit was placed inside the car with windows closed and tested. The Fig. 16 shows the decrease in the temperature of the Peltier device. This temperature reading was conducted at 10 am when the car was under the shade. The outside temperature was 35□C and the ambient temperature inside the car cabin is 39□C. The temperature is set at 23□C. Within 4-5 minutes the temperature reached 23□-degree Celsius and maintained.

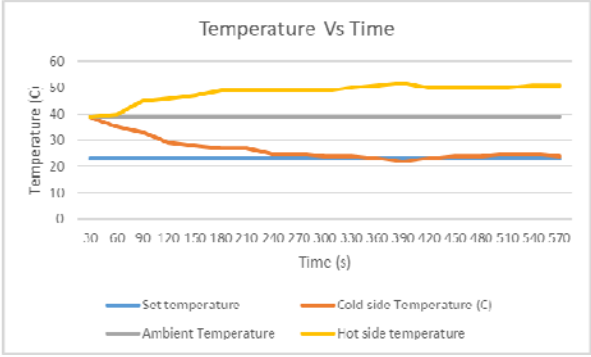


Fig. 16: Temperature vs graph of TEC

The open-loop testing was conducted to find PID parameters. The transient response was found in PSpice by constructing the PID circuit and TEC equivalent circuit. The steady-state gain Kc is seen as a proportion of the maximum open-loop response or the input voltage (delta y), which is set to 10, to step response (delta u), which is set to 6.

$$Kc = \frac{\Delta y}{\Delta u} = \frac{10}{6} = 1.66 \quad (3)$$

$$Td = 0.8ms$$
$$Ts = 1.98s.$$

The gain Kc, time delay Td, and starting time Ts were substituted in Zeigler Nicholas constants equation and the value is given in table 1.

TABLE I. CALCULATED PID PARAMETERS

Controller Type	Kp	Ti	Td
PID	4930	1.6ms	0.4ms

All three controllers were tested in MATLAB with the calculated parameters and found that the PI controller produced the best result due to the short settling time of 2s.

The PID coefficient was used to find the corresponding duty cycle. Substituting the new duty cycle in the input of the H-bridge gives a smoother graph in Fig. 17. As required, a voltage of 6V was produced at the terminal. Therefore, these PID parameters correspond to providing an appropriate output.

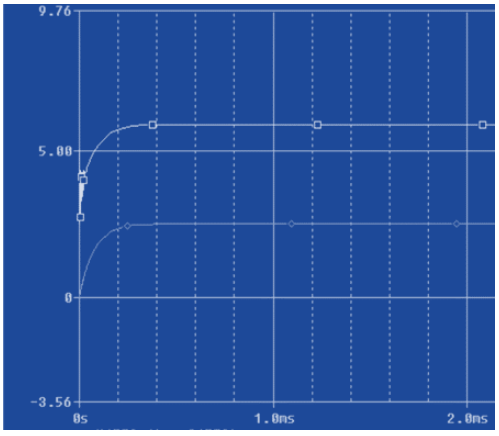


Fig. 17: Voltage and current are seen over one period of PWM. (Top graph: Voltage; Bottom Graph: Current)

VI. Conclusion

By utilizing the information collected from the literature review, a prototype is designed, fabricated, and tested. The health and safety of the child inside a car for a prolonged time are researched. Multiple types of the thermoelectric cooling system from several scholars were researched and found that by increasing the number of single-stage thermoelectric units use less power and provides better cooling. The heat flux inside the car cabin was researched and found that the surrounding heat load was increasing. The excess heat produced from the thermoelectric unit can be overcome using heatsink and forced convection such as a centrifugal fan. Moreover, the type of fabric such as mesh cotton fabric can greatly improve the thermal conduction. The thermoelectric device must be automatically controlled using PID control to efficiently manage the output temperature on the device.

Acknowledgment

My special thanks to Mr. Mohammed Al-Musleh for assisting me in developing this project.

References

[1] S. P. e. al., "Thermoelectric cooler and thermoelectric generator devices: A review of present and potential applications, modeling and materials.," *Energy,* vol. vol. 186., 2019.

[2] W. W. A. E. Gaurang Gupta, "Vehicle seat with integrated temperature control system". The United States Patent 0291613 A1, 26 September Sept 26, 2019.

[3] A. H. S. C. Salter, "Vehicle Having Moisture Sensing". The United States Patent 10464454, 5 November 2019.

[4] I.-V. M. Industries, "THERMOELECTRIC TECHNOLOGIES USED IN THE AUTOMOTIVE INDUSTRY," 17 January 2019. [Online]. Available: https://blog.marlow.com/thermoelectric-technologies-used-in-the-automotive-industry.

[5] T. Research, "Global Baby Car Seat Market 2017-2021: Market Research and Segmentation by Technavio.," Global Baby Car Seat Market 2017-2021: Market Research and Segmentation by Technavio, 2015. [Online]. Available: https://www.businesswire.com.

[6] A. M. a. M. Bouzidi, "Computation of thermal comfort inside a passenger car compartment," *Applied Thermal Engineering,* vol. vol. 26, no. no. 14, p. pp. 1697–1704, 2006.

[7] NCBI, "Temperature measurement in pediatrics," *Paediatrics & child health,* Vols. vol. 5, no. 5, pp. 273-84, 2000.

[8] A. Grundstein, "Quantifying the Heat-Related Hazard for Children in Motor Vehicles," *Bulletin of the American Meteorological Society,* 2010.

[9] B. Yu, "Smart Car Seat Helps Prevent Child Hot Car Deaths," KPIX5 SF BayArea, 2018. [Online]. Available: https://www.youtube.com/watch?v=qrPHokq7SIM.

[10] U. department, *The 2017 National Survey of the Use of Booster Seats,* US department of transportation, National Highway Traffic Safety Administration, 2017.

[11] Y. W. L. G. L. Z. J. L. T. M. Tingrui Gong, "Thermo-mechanical analysis on a compact thermoelectric cooler,", *Energy,* vol. Volume 172, pp. Pages 1211-1224, 2019.

[12] M. Zebarjadi, "Electronic cooling using thermoelectric devices"," *Appl. Phys.,* vol. Lett. 106.

[13] J. W. Finn, " An Integrated Child Safety Seat Cooling System—Model and Test.," *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY,* vol. VOL. 61, no. NO. 5, 2012.

[14] R. Roj, "Investigation on the Application of Different Air-Cooling-Systems in a Thermoelectric Setup," in *12th European Conference on Thermoelectrics*, 2015.

[15] " Optimization of Heat Sink Design and Fan Selection in Portable Electronics Environment.," 2001.

[16] N. Concepts, "Fan cooled heatsinks," [Online]. Available: https://www.novelconceptsinc.com/.

[17] A. Zaharia, "Centrifugal Fan CFD Analysis for Predicting its Aerodynamics," 11th March 2020.

[18] Y. W. P. Y. X. D. a. D. S. H. Du, "Experimental Investigation of a Temperature-Controlled Car Seat Powered by an Exhaust Thermoelectric Generator," *Journal of Electronic Materials,* vol. vol. 45, no. no. 3, p. pp. 1529–1539, 2016.

[19] A. V. W. Aniket A. Gabhane1, "Design of Comfortable Advanced Ventilated Automotive Seat for Driver using CFD simulation," *International Research Journal of Engineering and Technology (IRJET),* vol. Volume: 03, no. Issue: 07, 2016.

[20] A. Blades, "What Happens to Air Flow in Ducts When Size Changes," EnergyVGaurd, [Online]. Available: https://www.energyvanguard.com/blog/what-happens-air-flow-ducts-when-size-changes.

[21] A. Majumdar, "Modelling of thermal conductivity of knitted fabrics made of cotton-bamboo yarns using artificial neural network," *The Journal of The Textile Institute,* vol. vol. 102, no. no. 9, p. pp. 752–762, 2011.

[22] R. Senthilkumar, "Video Presentation of the project," [Online]. Available: https://youtu.be/etkeE5N4mEM.

[23] R. Senthilkumar, "Thermal Simulation of Thermoelectric cooling with infant model using Energy2D.," 2020.

Application of Nonlinear Estimation Strategies on a Magnetorheological Suspension System with Skyhook Control

Andrew S. Lee
Mechanical Engineering
University of Guelph
Guelph, Ontario, Canada
alee32@uoguelph.ca

S. Andrew Gadsden
Mechanical Engineering
University of Guelph
Guelph, Ontario, Canada
gadsden@uoguelph.ca

Mohammad Al-Shabi
Mechanical Engineering
University of Sharjah
Sharjah, UAE
malshabi@sharjah.ac.ae

***Abstract*—Extraction of state values from noisy or uncertain systems is important for feedback control because it improves the accuracy of the error signal. For known linear systems with Gaussian white noise, the Kalman filter provides optimal state estimates in terms of state error. However, electromechanical systems, such as magnetorheological dampers, typically exhibit nonlinear behaviour. In this paper, a new nonlinear estimation method known as the extended sliding innovation filter is presented and applied on a magnetorheological suspension system. The state estimates are extracted from a quarter car model with an active magnetorheological suspension system with skyhook control. The results are compared with the popular extended Kalman filter, and future experiments are considered.**

Keywords—estimation theory, actuator, control

I. Introduction

The objective of estimation theory is to extract information on states in the presence of system and measurement noise. For known linear systems with Gaussian white noise, the Kalman filter (KF) produces the optimal state estimate [1]. The Kalman gain minimizes the trace of the updated state error covariance matrix and is proven to produce the optimal state estimate for known linear systems with Gaussian noise [1, 2, 3]. The KF has a wide range of applications, in areas such as target tracking, signal processing, and fault detection [4, 5, 6]. Similar to control theory, a trade-off exists between estimation accuracy and robustness to disturbances. Current research in estimation theory aims to reduce state estimation error while simultaneously being robust to disturbances, faults, and modeling uncertainty [1].

For nonlinear systems, several strategies have been proposed. The extended Kalman Filter (EKF) can be used to estimate the states of a nonlinear dynamic system. The filter uses local linearization of the system model at the operating point in order to calculate the corrective Kalman gain [1, 4]. The EKF derivation is based on the first-order Taylor series expansion of the nonlinear system model, also known as a Jacobian matrix [1]. However, if the system is highly nonlinear, the EKF solution may diverge from the true state trajectory leading to numerical instabilities and poor estimation results [7].

Another nonlinear estimation strategy is the sigma-point Kalman filter (SPKF), which is formulated using a weighted statistical linear regression approach that linearizes the nonlinear state model statistically [8]. The SPKF produces sigma points from the projected probability distribution of states and are mapped through the nonlinear system model [8]. Unlike the EKF, this system does not require local linearization of the system model which improves the estimation process for nonlinear systems [4].

Similarly, the unscented Kalman filter (UKF) also uses sigma points from the projected probability distribution of states [8]. The UKF uses a deterministic sampling approach knows as the unscented transform to select a minimal set of sample points around the mean [1, 4]. It is possible to approximately determine the mean and covariance of the density using Monte Carlo sampling [1]. The UKF can capture the updated statistical mean and covariance up to the third-order for any nonlinearity; making it superior to the EKF in terms of estimation accuracy [1, 7]. However, like many particle filters, the UKF is resource intensive and may suffer from poor sampling rates [4].

The smooth variable structure filter (SVSF) was introduced in an effort to improve upon the robustness of estimation strategies [9]. The derivation of the SVSF gain yields a switching function based on variable structure techniques [9]. The gain is a function of measurement error and a switching term that keeps the estimates along the true state trajectory [9]. Stability is improved by defining an upper bound on noise and modeling uncertainty which improves robustness to uncertainties and disturbances.

Most recently, a new filter was introduced based on variable structure techniques similar to the SVSF called the sliding innovation filter (SIF) [10]. The SIF is a robust estimation strategy that provides a stable estimation process, and has been found to be more accurate than the SVSF [10]. Similar to the SVSF, the SIF creates an upper bound on noise and modeling uncertainty but uses a different method for calculating the corrective gain term. In this paper, similar to the EKF, the SIF is extended to nonlinear systems, and is referred to as the extended sliding innovation filter (ESIF). The ESIF is compared with the popular EKF as a benchmark. The filters are applied to an electromechanical system consisting of a quarter car model with skyhook control using a magnetorheological (MR) damper for active damping.

The paper is organized as follows. In Section II, the background on magnetorheological systems is provided. The ESIF equations are provided in Section III. The simulation setup and results are discussed in Section IV, followed by concluding remarks.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

II. Magnetorheological Systems

A. Magnetorheological Fluids

Magnetorheological fluids are a class of smart fluids that significantly change viscosity when a magnetic field is applied. The fluid consists of micron scale (1 to 10 μm) magnetically polarizable particles suspended in a carrier medium such as mineral or silicon oil [11]. Surfactants are often used in the non-colloidal mixture in order to prevent the settling of the suspended particles [12]. The magnetic particles typically account for 20 to 60 percent of the MR fluid's total volume [11]. The MR fluid reacts nearly instantaneously allowing for smooth force adjusting and repeatability [13].

When a magnetic field is applied to the MR fluid, the ferromagnetic particles rearrange to form linear structures and chains. The formation of these structures results in a change in the fluid's viscosity. Increasing the strength of the magnetic field increases the yield strength of the MR fluid [14]. Since the discovery of the MR fluid effect in the 1940s by Jacob Rabinov, commercial products have been developed to attain low viscosities and high yield stresses; a 100 kPa yield stress is achievable within a few milliseconds [14, 15].

This property has three main engineering application modes: flow (valve) mode, shear mode and squeeze mode [12]. In flow mode, the magnetic field is normal to the flow of the MR fluid and is typically used for linear dampers [12]. In shear mode, the magnetic field is normal to the displacement of the shear walls and is used for rotary dampers, breaks, and clutches [12]. Finally, squeeze mode utilizes a magnetic field that is parallel to the desired displacement of the containing walls. The squeeze mode provides large forces for relatively small displacements [12].

B. Magnetorheological Dampers

The quick response time of MR fluid behavior when exposed to a magnetic field is ideal for electromechanical devices such as actuators and dampers. Semi-active MR control devices have the versatility of active dampers while retaining the reliability of passive ones [12]. The two main engineering applications of MR fluids are linear and rotary dampers.

The main advantage of linear MR dampers is the controllability of the system damping by changing the input current. By changing the current, the magnetic induction in an orifice between two separated MR fluid chambers is regulated [12]. Linear MR dampers such as the one shown in Fig. 1 use the MR fluid in flow mode to treat the orifice as a valve for the MR fluid. One of the first applications of linear MR dampers in the automotive industry is its use as a secondary suspension element for on and off highway vehicles [15]. By controlling the damping arrangement, the excitation frequency and vibration transmission can be adjusted as desired [15].

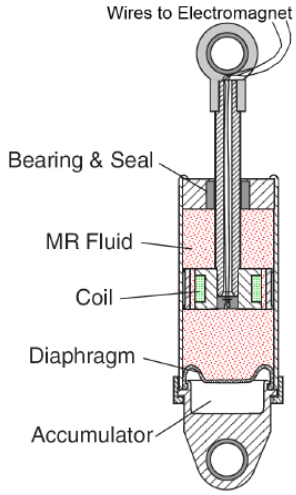


Figure 1. Schematic of a single ended linear MR damper [12].

C. Magnetorheological Force-Velocity Hysteresis Model

Several different models have been proposed in literature to describe the force-velocity hysteresis curve of linear MR dampers including bi-viscous, polynomial, Bouc-Wen, and Sigmoid [14, 15, 16]. While the Bouc-Wen produces the best estimates according to [16], the polynomial model produces reasonable estimates and was chosen due to ease of implementation for the EKF and proposed ESIF.

The force-velocity relationship of an MR damper is highly dependent on the applied current. The damping force f_d of the MR damper can be described by the following:

$$f_d = f_i(i) f_h(y, \dot{y}, \ddot{y}) \tag{2.1}$$

where f_i is a function of current i, and f_h describes the hysteretic behavior of the piston displacement, velocity, and acceleration denoted by y, $\dot{y}$, and $\ddot{y}$, respectively [16]. The nonlinear incremental current proposed by [16] is as follows:

$$f_i(i) = 1 + \frac{k_i}{1 + e^{-a(i+I_0)}} - \frac{k_i}{1 + e^{-a \cdot I_0}};\ i \geq 0 \tag{2.2}$$

where k, a, and I_0 are constants obtained through measured data. A model of the piecewise polynomial function to describe the MR damper during positive and negative acceleration was first proposed by [17]. The polynomial hysteresis function is as follows:

$$f_h = \left(\sum_{k=0}^{n} b_k \dot{y}\,(t)^k \right);\ n = 6 \tag{2.3}$$

where b_k represents the polynomial coefficient constants obtained through experimentation, k is the polynomial exponent, and n is the order of the polynomial. By combining (2.2) and (2.3) the damping force function becomes:

$$f_d = \begin{cases} f_i(i) \left(\sum_{k=0}^{6} b_{uk} \dot{y}(t)^k \right);\ \ddot{x} < 0 \\ f_i(i) \left(\sum_{k=0}^{6} b_{dk} \dot{y}(t)^k \right);\ \ddot{x} > 0 \\ f_i(i) \left(\sum_{k=0}^{6} \frac{1}{2}(b_{uk} + b_{dk}) \dot{y}(t)^k \right);\ \ddot{x} = 0 \end{cases} \tag{2.4}$$

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

where b_{uk} and b_{dk} represent the coefficients of the upper and lower polynomial, respectively [16]. In order to ensure the convergence of the two polynomial functions near the extremities, when the piston acceleration is zero, the damping force is calculated as the average of the upper and lower polynomial functions [16]. The value of the constants and polynomial coefficients used in this paper can be found in the Appendix.

III. The Extended Sliding Innovation Filter

In this paper, a nonlinear form of the SIF is presented referred to as the extended sliding innovation filter (ESIF). Similar to the extended Kalman filter (EKF), the proposed ESIF makes use of the linearized form of the nonlinear system and measurement functions [4]. For example, consider the nonlinear system function $f(\hat{x}_{k|k}, u_k)$ and the nonlinear measurement function $h(\hat{x}_{k+1|k})$. Linearized forms of these nonlinearities may be calculated using the following partial derivatives (Jacobians), respectively:

$$F_{k+1} = \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_{k|k}, u_k} \tag{3.1}$$

$$H_{k+1} = \left.\frac{\partial h}{\partial x}\right|_{\hat{x}_{k+1|k}} \tag{3.2}$$

where x refers to the state, $\hat{x}$ refers to the state estimate, u refers to the control signal input, k refers to the time step, $k+1|k$ refers to the prediction step, and $k|k$ refers to the previous time update step. The structure of the nonlinear SIF estimation process is similar to the EKF with the main difference being the formulation of the gain. The SIF and ESIF are both predictor-corrector estimators. The prediction stage consists of three main equations, as follows:

$$\hat{x}_{k+1|k} = f(\hat{x}_{k|k}, u_k) \tag{3.3}$$

$$P_{k+1|k} = F_{k+1} P_{k|k} F_{k+1}^T + Q_{k+1} \tag{3.4}$$

$$\tilde{z}_{k+1|k} = z_{k+1} - h(\hat{x}_{k+1|k}) \tag{3.5}$$

Note that f refers to the nonlinear system function, F_k refers to the linearized version of the system matrix A (Jacobian matrix or first-order Taylor series expansion) at time k, and h refers to the nonlinear measurement function. The states are first predicted in (3.3) before being updated in (3.7) using the gain defined in (3.6). The state error covariance matrix is first predicted in (3.4) before being updated in (3.8). Note that the gain (3.6) is also used to update the state error covariance (3.8). The update stage consists of three main equations, as follows:

$$K_{k+1} = H_{k+1}^{+} \overline{sat}\left(\left|\tilde{z}_{k+1|k}\right|/\delta\right) \tag{3.6}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\tilde{z}_{k+1|k} \tag{3.7}$$

$$P_{k+1|k+1} = (I - K_{k+1}H_{k+1})P_{k+1|k}(I - K_{k+1}H_{k+1})^T \ldots \\ \ldots + K_{k+1}R_{k+1}K_{k+1}^T \tag{3.8}$$

Note that H_{k+1}^{+} refers to the pseudoinverse of the linearized measurement matrix (first-order Taylor series expansion) at time k, $\overline{sat}$ refers to the diagonal of the saturation term, sat refers to the saturation of a value (yields a result between -1 and +1), $\left|\tilde{z}_{k+1|k}\right|$ refers to the absolute value of the innovation, δ refers to the sliding boundary layer, I refers to the identity matrix (of dimension n-by-n where n is the number of states), R refers to the measurement noise covariance, and $k+1|k+1$ refers to the updated values at time $k+1$ based on information at time $k+1$. Equations (3.1) through (3.8) represent the proposed ESIF estimation process for nonlinear systems and measurements.

IV. Computer Simulations and Results

A. Quarter Car Model with Active Damper

The EKF and ESIF were applied to a quarter car model based on Fig. 2. The equations of motion are defined as follows:

$$M_s\ddot{x}_1 + k_s(x_1 - x_3) + f_d = 0 \tag{4.1}$$

$$M_u\ddot{x}_3 + k_s(x_3 - x_1) + k_t(x_3 - r) + \cdots \\ \cdots b_s(\dot{x}_3 - \dot{x}_1) + b_t(\dot{x}_3 - \dot{r}) + f_d = 0 \tag{4.2}$$

where M_s and M_u are the sprung mass and unsprung mass, respectively, k_s and b_s are the spring constant and damping coefficient between the sprung mass and the unsprung mass, respectively, k_t and b_t are the spring constant and damping coefficient of the tire, respectively, and r is the road profile (defined later in Fig. 3). As stated previously, f_d is the damping force exerted by the MR damper. The state space equations based on this model and used in the simulation are defined as follows:

$$\dot{x}_1 = x_2 \tag{4.3}$$

$$\dot{x}_2 = (k_s(x_3 - x_1) - f_d)/M_s \tag{4.4}$$

$$\dot{x}_3 = x_4 \tag{4.5}$$

$$\dot{x}_3 = (k_s(x_1 - x_3) + k_t(x_1 - r_3) + \cdots \\ \cdots b_s(x_2 - x_4) + b_t(\dot{r} - x_4) - f_d)/M_u \tag{4.6}$$

The value of the constants used in the simulation are available in the Appendix. The skyhook criterion used to control the MR damper is described as follows:

$$x_2(x_2 - x_4) \geq 0;\ \ i = 1\ amp$$
$$x_2(x_2 - x_4) = 0;\ i = 0\ amp \tag{4.7}$$

If the relative velocity of the sprung mass with respect to the unsprung mass is in the same direction as the velocity of the unsprung mass, then a current is applied to the MR damper in order to reduce the body acceleration [18]. For this controller, estimation of the velocities states of each mass is crucial in order to make sure the system meets the skyhook criterion for MR damper activation.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

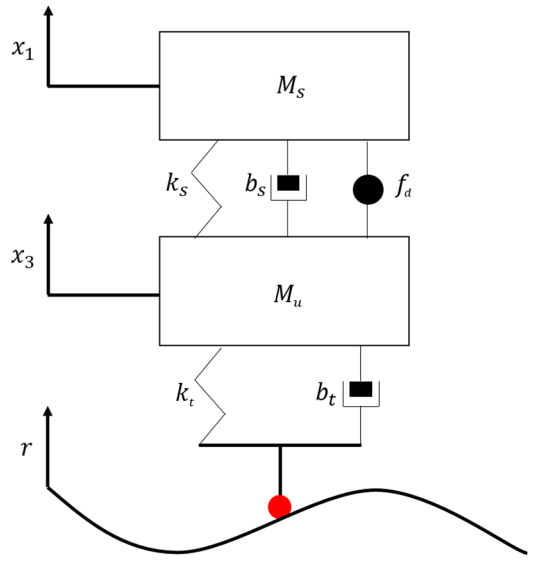


Figure 2. Quarter car model used in simulation to compare EKF and ESIF.

The EKF and ESIF methods were investigated using Matlab to simulate the quarter car model, implement a skyhook controller, and generate a road profile. The simulation used a flat road profile with a bump encountered at four seconds (a disturbance) and a sample rate of 1 ms. The bump was simulated using a sinusoid with an amplitude of 5 cm and frequency of 0.5 Hz as shown in Fig 3.

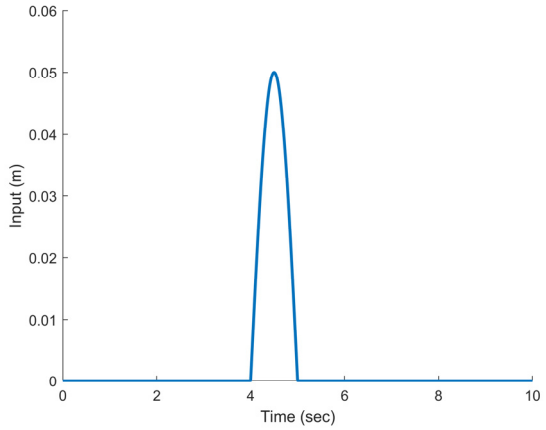


Figure 3. Road profile (with a disturbance) used in simulation.

The quarter car model was first tested without the activation of the skyhook controller and MR damper in order to visualize the free dynamics of the system. The position of the sprung mass is shown in Fig. 4. Note that the system and measurement noise covariances used in the simulation were defined as Q and R, respectively:

$$Q = diag[10^{-9} \quad 10^{-8} \quad 10^{-9} \quad 10^{-8}] \qquad (4.8)$$

$$R = diag[10^{-5} \quad 10^{-4} \quad 10^{-5} \quad 10^{-4}] \qquad (4.9)$$

The EKF and SIF have similar performances when predicting the states when no controller or MR damper is applied. This is because without the use of the MR damper, the quarter car system is linear and uses Gaussian white noise to describe the system and measure noise. In this scenario, the EKF should provide similar performance to the ESIF.

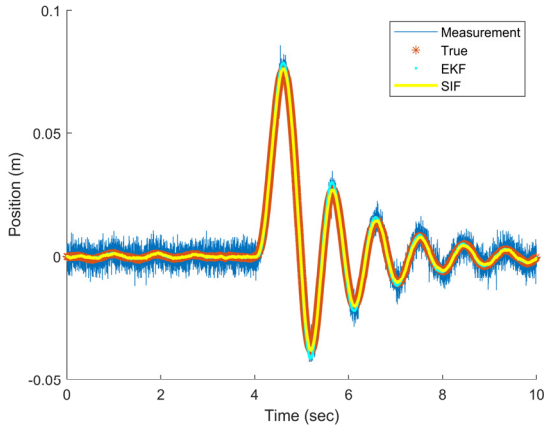


Figure 4. Sprung mass position without skyhook controller.

The skyhook controller was implanted by using velocities state estimates from the EKF and ESIF as feedback. The sliding boundary layer widths used for the ESIF were tuned manually based on reducing estimation error, and defined as:

$$\delta = diag([0.5 \quad 1 \quad 0.5 \quad 0.09]) \qquad (4.10)$$

The position of the sprung mass using the EKF and ESIF in the skyhook controller is shown in Fig. 5.

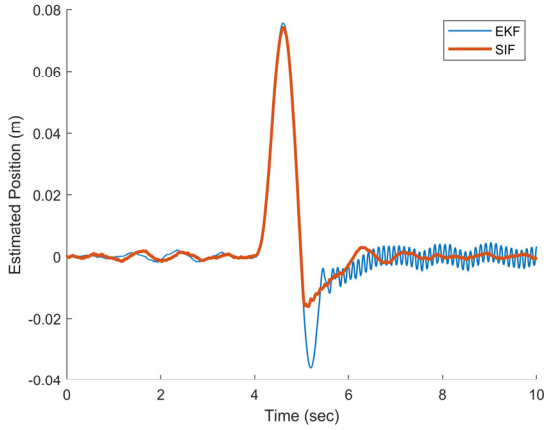


Figure 5. Skyhook controler performance using EKF and ESIF.

Both filters perform comparatively before the MR damper and skyhook controller is activated (at approximately five seconds) because the road profile is flat and the system dynamics remain relatively linear before the MR damper is activated. The initial bumps are a result of system noise.

When the skyhook controller is activated, the system dynamics become highly nonlinear due to the piecewise 6th-order polynomial force-velocity hysteresis model of the MR damper. This results in poor velocity state estimates for the EKF. The local linearization of the system cannot accurately model the system. The poor velocity state estimates result in erratic behavior of the skyhook controller as seen by the chattering in Fig. 5. The velocity estimates of the EKF can be seen in Fig. 6. The EKF reports higher velocities than the true values and does not capture the magnitude of the oscillations.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

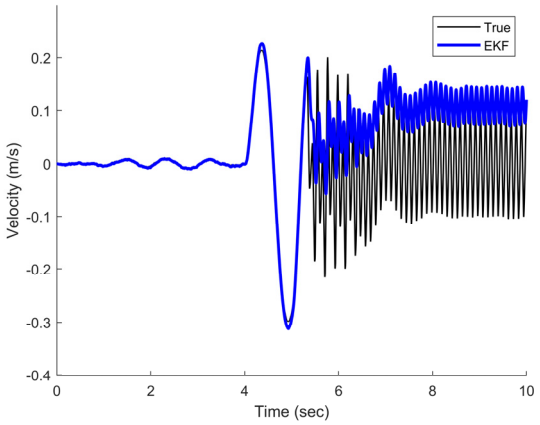


Figure 6. EKF sprung mass velocity estimates.

In order to fairly compare the EKF with the ESIF, the accuracy of both filters was assessed using the same controller input for the predicted state estimate. The result of the simulation is shown in Fig. 7. During the initial activation of the MR damper, the EKF deviates from the true state value as seen at five seconds. The EKF is able to converge as the amplitude of the sprung mass oscillation decreases, however it does so more slowly when compared with the ESIF.

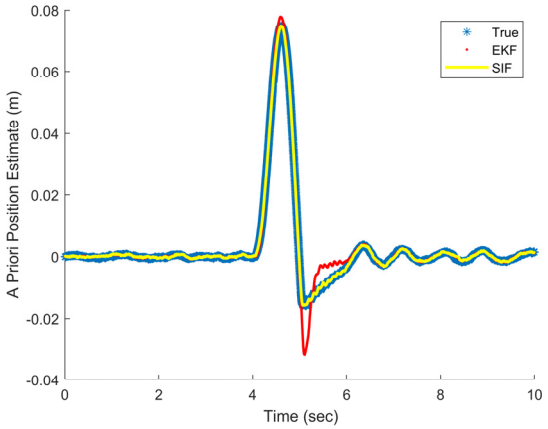


Figure 7. Comparison of predicted state estimates for EKF and ESIF.

The normalized error squared (NES) was plotted for both filters and is shown in Fig. 8. The EKF estimates can be shown to deviate from the true states when the MR damper is activated as shown by the sharp error spikes around four seconds, while the ESIF estimates follow the true states relatively closely throughout the entire simulation. This further highlights the robustness of the ESIF due to the switching function of the gain.

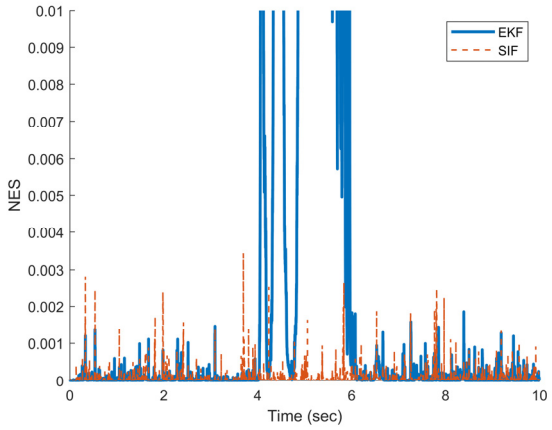


Figure 8. Normalized error squared (truncated at 0.01).

As shown in Table I, the root mean square error (RMSE) was calculated for each filter [4]. The ESIF was found to outperform the EKF, especially in terms of velocity state estimation, which is a crucial component of the skyhook controller.

TABLE I. RMSE VALUES

Filter	State			
	Sprung Mass Position (m)	***Sprung Mass Velocity (m/s)***	***Unsprung Mass Position (m)***	***Unsprung Mass Velocity (m/s)***
EKF	0.0023	0.0154	0.0008	0.0121
ESIF	0.0003	0.0013	0.0002	0.0043

V. CONCLUSIONS

While extracting state values from linear systems using model-based filters is relatively simple to implement, estimating states from nonlinear systems are far more challenging. Highly nonlinear systems such as MR dampers require novel approaches to achieve robust and accurate estimation. This paper presented a novel nonlinear estimation filter called the extended sliding innovation filter (ESIF). The ESIF uses a sliding boundary layer to represent the upper bounds of noise and modeling uncertainty. Using this approach, the ESIF was able to outperform the EKF which is known to have difficulties with highly nonlinear systems and uncertainties. The ESIF improved performance of the skyhook controller which was highly dependent on the accuracy of the velocity state estimates of the system.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

APPENDIX

In this section, the parameters used in the simulation for the hysteresis and quarter car models are provided.

TABLE II. PARAMETERS FOR HYSTERESIS MODEL

Parameter	Value
k_i	3.10
a	18.53
I_0	-0.15
b_{u0}	0.06
b_{u1}	4.97
b_{u2}	-71.29
b_{u3}	$-1{,}461.82$
b_{u4}	23,232.45
b_{u5}	190,972.68
b_{u6}	$-2{,}319{,}223.80$
b_{d0}	-0.07
b_{d1}	5.03
b_{d2}	78.00
b_{d3}	-1549.09
b_{d4}	$-27{,}398.44$
b_{d5}	210,738.82
b_{d6}	3,017,864.40

TABLE III. PARAMETERS FOR QUARTER CAR MODEL

Parameter	Value
M_s	$290\ Kg$
M_u	$59\ Kg$
k_s	$14{,}500\ N/m$
b_s	$1{,}385.4\ N \cdot s/m$
k_t	$19{,}000\ N/m$
b_t	$170\ N \cdot s/m$

REFERENCES

[1] H. H. Afshari, S. A. Gadsden and S. R. Habibi, "Gaussian Filters for Parameter and State Estimation: A General Review and Recent Trends," *Signal Processing,* vol. 135, pp. 218-238, 2017.

[2] B. Ristic, S. Arulampalam and N. Gordon, Beyond the Kalman Filter: Particle Filters for Tracking Applications, Boston: Artech House, 2004.

[3] S. Haykin, Kalman Filtering and Neural Networks, New York: John Wiley and Sons, Inc., 2001.

[4] S. A. Gadsden, "Smooth Variable Structure Filtering: Theory and Applications," McMaster University, Hamilton, 2011.

[5] S. A. Gadsden, M. Al-Shabi and S. R. Habibi, "Estimation Strategies for the Condition Monitoring of a Battery System in a Hybrid Electric Vehicle," *ISRN Signal Processing,* 2011.

[6] S. A. Gadsden, Y. Song and S. R. Habibi, "Novel Model-Based Estimators for the Purposes of Fault Detection and Diagnosis," *IEEE/ASME Transactions on Mechatronics,* vol. 18, no. 4, pp. 1237-1249, 2013.

[7] S. A. Gadsden, S. R. Habibi and T. Kirubarajan, "Kalman and Smooth Variable Structure Filters for Robust Estimation," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 50, no. 2, pp. 1038-1050, 2014.

[8] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE,* vol. 92, no. 3, 2004.

[9] S. R. Habibi, "The Smooth Variable Structure Filter," *Proceedings of IEEE,* vol. 95, no. 5, pp. 1026-1059, 2007.

[10] S. A. Gadsden and M. Al-Shabi, "The Sliding Innovation Filter," *IEEE Access,* vol. 8, pp. 96129-96138, 2020.

[11] A. Milecki, "Investigation and Control of Magneto-Rheological Fluid Dampers," *International Journal of Machine Tools and Manufacture,* vol. 41, no. 3, pp. 379-391, 2001.

[12] A. Spaggiari, "Properties and Applications of Magnetorheological Fluids," *Frattura ed Integrità Strutturale,* vol. 23, no. 23, pp. 57-61, 2013.

[13] D. Guth and J. Mas, "Energy-Efficient Clutch with Optimized Torque Density," in *Smart Materials, Adaptive Structures and Intelligent Systemd*, Snowbird, Utah, 2013.

[14] P. Yadmellat and M. R. Kermani, "Adaptive Hysteresis Compensation for a Magneto-Rheological Robot," in *IEEE Intelligent Robots and Systems (IROS)*, Tokyo, Japan, 2013.

[15] A. Olabi and A. Grunwald, "Design and Application of Magneto-Rheological Fluid," *Materials and Design,* vol. 28, no. 10, pp. 2658-2664, 2007.

[16] X. Q. Ma, S. Rakesha and C. Su, "Development and Relative Assessments of Models for Characterizing the Current Dependent Hysteresis Properties of Magnetorheological Fluid Dampers," *Journal of Intelligent Material Systems and Structures,* vol. 18, pp. 487-502, 2007.

[17] S. B. Choi, B. K. Lee and Y. P. Park, "A Hysteresis Model for the Field-dependent Damping Force of a Magnetorheological Damper," *Journal of Sound and Vibration,* vol. 2, no. 245, p. 375, 2001.

[18] H. Sleiman, B. Lemaire-Semail, B. Clenet and J. Lozada, "Modeling and Inversion-based Control of a Magnetorheological Vehicle Suspension," in *IEEE Vehicle Power and Propulsion Conference*, Lille, France, 2010.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

A Study of Variable Structure and Sliding Mode Filters for Robust Estimation of Mechatronic Systems

S. Andrew Gadsden
Mechanical Engineering
University of Guelph
Guelph, Ontario, Canada
gadsden@uoguelph.ca

Mohammad Al-Shabi
Mechanical Engineering
University of Sharjah
Sharjah, UAE
malshabi@sharjah.ac.ae

***Abstract*—In this paper, a study of estimation strategies based on variable structure and sliding mode theory is performed. The smooth variable structure filter (SVSF) and the new sliding innovation filter (SIF) are based on similar sliding mode concepts but with some notable differences. The relevant literature and background are explored and the SVSF and SIF estimation algorithms are presented. For comparison purposes, the two estimation strategies are applied on a mechatronic system. The results indicate that although both the SVSF and SIF provide robust estimates to faults, the SIF formulation provides slightly more accurate estimates while maintaining robustness, and is less computationally complex.**

Keywords—estimation theory, robustness, sliding innovation filter, smooth variable structure filter, variable structure

I. Introduction

Estimation theory is typically found in the field of signal processing and involves the use of filters to extract useful system information from noisy measurements [1]. For example, consider a mechatronic system found in aerospace such as a flight surface actuator. One of the most important states used for successful maneuvers of aircraft is the angle of the flap. A sensor is used to provide measurements of the angle; however, sensors are usually corrupted by unwanted signals such as measurement noise. It is the goal of an estimation method, such as the popular and well-known Kalman filter (KF), to extract the true state value from the measurement noise [2]. This value would then be used by the flight control computer to adjust control signals to ensure desired flight performance. The estimation method would also need to minimize the effects of system noise (e.g., natural flight surface vibrations or internal system vibrations).

Most estimation strategies found in the literature are based on statistical methods and Bayesian theory [1]. The KF offers the optimal solution to the linear estimation problem as it minimizes the state error covariance, which is a second-order moment that provides a means to approximate the state error. The state error is defined as the difference between the true state value and the estimated value. Although the KF is considered an optimal filter for linear systems, it is derived based on a number of strict assumptions such as the noise must be white (or zero mean and with Gaussian distributions), and the system and measurement dynamics must be known [3]. For nonlinear systems and measurements, a number of different forms of the KF have been presented. The two most common are the extended Kalman filter (EKF) which uses first-order Taylor series approximations, and the unscented Kalman filter (UKF), which uses the mathematical unscented transform to approximate the nonlinearities. The UKF is a type of sigma-point filter as a number of points are derived and propagated through the nonlinearities to provide a higher-order approximation when compared with the EKF [3]. The EKF and UKF methods are considered sub-optimal solutions to the nonlinear estimation problem as there is currently no known analytical solution [4, 5].

Although popular, KF-based methods suffer from stability and robustness issues [1]. For example, should the system model being used by the filter derive from the true system dynamics, the estimation results will deviate from the truth. This proves a challenge for sensitive control systems to ensure safe and reliable operation. A number of improvements have been suggested in literature to improve the robustness and stability of KF-based methods [6, 7, 8]. Some methods have looked at improving numerical stability through the use of square-root formulations [9, 10]. Others have looked at bounding the estimates and covariance values to ensure the estimates remain close to the true state values [11]. Some literature explored the use of modifying the system noise covariance or adding some degree of memory [4, 12]. For estimation problems which contain a lot of modeling uncertainty, the system noise covariance can be made artificially large to capture a wider-range of values [4, 12]. This has led to dual-estimation strategies where two KFs are run in parallel and small and large system noise covariances are implemented, respectively. The measurement error, also known as the innovation, could be used to determine which KF yields the better estimate. This strategy is considered a basic adaptive KF approach [13].

Sliding mode observers and variable structure approaches were introduced to improve the robustness of the estimation process [12]. These methods make use of a variable structure gain which essentially implements a boundary layer across the true state trajectory and switches back-and-forth across it. Although there are a number of other filters and strategies based on sliding mode and variable structure theory, this study will focus on the smooth variable structure filter (SVSF) and the sliding innovation filter (SIF) as they are derived using the same principles. The SVSF was introduced in 2007, and was

978-1-7281-9615-2/20/$31.00

formulated as a predictor-corrector filter based on sliding mode and variable structure techniques [14]. The estimation process for the SVSF is similar to the KF, except that it is considered a sub-optimal filter since its gain was derived for robustness as opposed to optimality [14]. The SVSF gain is a function of measurement errors and a switching term. This gain formulation allowed for a robust estimation strategy which reduced the negative effect of modeling uncertainties and disturbances.

The SVSF presented in [14] did not contain a state error covariance function which provides a measure of performance as well as can be used to expand the number of useful applications. In [2, 15], a new formulation of the SVSF was presented which included the derivation for a state error covariance function. This increased the number of useful applications, and allowed for the SVSF to be combined with the interacting multiple model (IMM) method for target tracking and fault detection and diagnosis purposes [16]. The proposed fault detection strategy, the SVSF-IMM, yielded higher detection rates for faults and minimized misclassifications of faults when compared with the KF-IMM. In [17], the chattering effects of the SVSF gain were used to detect faults. Since the SVSF gain is a function of the measurement error, this process was essentially the same as using the magnitudes of the innovation as others have proposed in literature and practice [1]. Another formulation of the SVSF for fault detection and diagnosis was presented in [18], and was found to yield good fault detection accuracy.

A higher-order form of the SVSF was proposed in [19]. The results of the proposed second-order SVSF indicated that it was more accurate than the standard SVSF, however it was more computationally expensive and complex. Another type of higher-order SVSF was presented in [20]. Due to the state error covariance derivation of the SVSF presented in [2, 15], the SVSF was reformulated for a number of different applications such as object target tracking [13, 21, 22]. The SVSF has also been applied to other robotics [23, 24], mechatronics [25], and aerospace estimation problems [26, 27, 28, 29] and the literature continues to expand.

The SIF is a new filter introduced in 2020, and is based on similar concepts to the SVSF [30]. Like the SVSF, it can be applied to both linear and nonlinear systems and measurements. However, the SIF utilizes a simpler gain structure which reduces the computational complexity [30]. This allows for faster online implementations and real-time embedding within sensors and chips. The results presented in [30] did not compare the new SIF with the SVSF.

In this paper, the SVSF and SIF are applied on a linear mechatronic system, and the results are compared. Since the SIF is closely related to similar concepts as the SVSF, their differences are important to share and discuss as it adds to the overall body of literature. The paper is organized as follows. For completeness, the SVSF equations are provided in Section 2, followed by the SIF equations in Section 3. The simulation results are provided and discussed in Section 4, followed by concluding remarks.

II. The Smooth Variable Structure Filter

In this section, the linear smooth variable structure filter (SVSF) estimation process with a covariance derivation is summarized. The SVSF may also be formulated to handle nonlinear systems and measurements. For this implementation, it is assumed that the linear systems and measurements are modelled using the following respective state space equations:

$$x_{k+1} = Ax_k + Bu_k + w_k \tag{2.1}$$

$$z_{k+1} = Cx_{k+1} + v_{k+1} \tag{2.2}$$

where x refers to the states, z refers to the measurements, u refers to the system input, A refers to the system matrix, B refers to the input gain matrix, C refers to the measurement matrix, and k refers to the time step. The system and measurement noises are defined using w and v, respectively. Furthermore, the system and measurement noise are assumed to be zero mean with Gaussian covariance's Q and R, respectively.

The SVSF is formulated as a predictor-corrector estimator. During the prediction stage, the state estimates and state error covariances are predicted, as well as the predicted measurement error, respectively as follows:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_k \tag{2.3}$$

$$P_{k+1|k} = AP_{k|k}A^T + Q_{k+1} \tag{2.4}$$

$$\tilde{z}_{k+1|k} = z_{k+1} - C\hat{x}_{k+1|k} \tag{2.5}$$

where the subscript $k|k$ refers to the updated value at the previous time step, and the subscript $k+1|k$ refers to the predicted value at the current time step.

The SVSF update stage includes the calculation of the SVSF gain, the updated state estimates, updated state error covariance, and the updated measurement error. These are respectively calculated using the following equations:

$$K_{k+1} = C^+\left(\left|\tilde{z}_{k+1|k}\right| + \gamma\left|\tilde{z}_{k|k}\right|\right) \circ sat\left(\tilde{z}_{k+1|k}/\psi\right) \tag{2.6}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\tilde{z}_{k+1|k} \tag{2.7}$$

$$P_{k+1|k+1} = (I - K_{k+1}C)P_{k+1|k}(I - K_{k+1}C)^T \ldots \\ \ldots + K_{k+1}R_{k+1}K_{k+1}^T \tag{2.8}$$

$$\tilde{z}_{k+1|k+1} = z_{k+1} - C\hat{x}_{k+1|k+1} \tag{2.9}$$

where $+$ refers to the pseudoinverse, γ represents the SVSF 'memory' and is a value between 0 and 1, $\circ$ is the Schur product (element by element multiplication), sat refers to the saturation function (returns values between -1 and $+1$), ψ is the SVSF smoothing boundary layer width (and has a value for each measurement), and I refers to the identity matrix (of dimension n-by-n where n is the number of states).

Equations (2.3) through (2.9) inclusively represent the SVSF estimation process for linear systems and measurements, and is repeated iteratively. Note that the process requires initialized state estimates and state error covariances. Note also that $\tilde{z}_{k|k}$ found in (2.6) refers to (2.9) from the previous time step.

The smoothing boundary layer ψ refers to the amount of uncertainties present in the estimation process. The greater the amount of uncertainties, such as noise, modeling uncertainties, and disturbances, the larger its defined value. The value, or width, may be set as a fixed value or can be made time-varying. Note also that if the boundary layer width is chosen to be too small, chattering or high-frequency switching may occur due to the uncertainties being underestimated.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

III. The Sliding Innovation Filter

In this section, the sliding innovation filter (SIF) estimation process is summarized. Note that the SIF is also formulated as a predictor-corrector estimator, and can be applied to both linear and nonlinear systems. However, since a linear mechatronic system is considered in Section IV, the linear SIF estimation process will be summarized. The SIF estimation concept is visualized in Fig. 1.

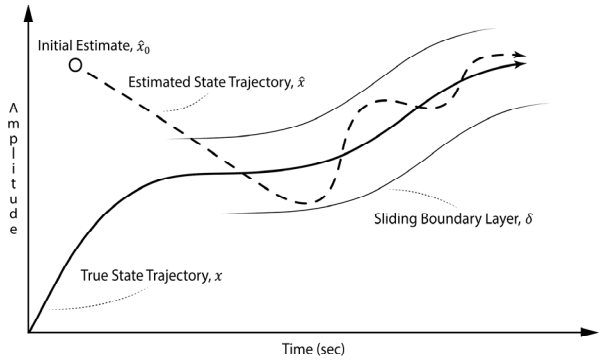


Figure 1. The SIF estimation process is summarized here [30]. The initial estimates are forced to converg along the true state trajectory, and switches back and forth due to the switching effect in the SIF gain. The values are bounded to within a region of the true trajectory by the sliding boundary layer width which is defined based on the amount of uncertainties.

Similar to the SVSF, the prediction stage includes calculating the predicted state estimates, the predicted state error covariance, and the predicted innovation, as per the following three equations, respectively:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_k \tag{3.1}$$

$$P_{k+1|k} = AP_{k|k}A^T + Q_k \tag{3.2}$$

$$\tilde{z}_{k+1|k} = z_{k+1} - C\hat{x}_{k+1|k} \tag{3.3}$$

The update stage includes calculating the SIF gain, the updated state estimates, and the updated state error covariance, as per the following three equations, respectively:

$$K_{k+1} = C^{+}\overline{sat}\left(\left|\tilde{z}_{k+1|k}\right|/\delta\right) \tag{3.4}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\tilde{z}_{k+1|k} \tag{3.5}$$

$$P_{k+1|k+1} = (I - K_{k+1}C)P_{k+1|k}(I - K_{k+1}C)^T \ldots + \ldots + K_{k+1}R_{k+1}K_{k+1}^T \tag{3.6}$$

where $\overline{sat}$ refers to the diagonal of the saturation term, and δ refers to the sliding boundary layer width. Equations (3.1) through (3.6) represent the SIF estimation process for linear systems and measurements defined by (2.1) and (2.2), respectively.

The notable difference between the SVSF and SIF estimation strategies are the gains defined by (2.6) and (3.4). The SVSF gain as defined in (2.6) requires two tuning parameters (γ and ψ), a Schur product calculation, and the updated measurement error from the previous time step. However, the SIF gain as defined in (3.4) requires only one tuning parameter (δ), and there are no values required from the previous time step. The SVSF and SIF gains are inherently robust to modeling uncertainties and disturbances due to the switching effects of the saturation function. However, the SIF gain is simpler and requires less computational power.

Finally, note that both the SVSF and SIF have stability proofs defined by Lyapunov theory [14, 30]. Their respective proofs of stability show that the measurement error must decrease with time for the methods to be considered stable. Both filters have their gains derived based on the proofs of stability, which allows for robust estimation processes.

IV. Computer Experiments and Results

In this section, the KF, SVSF, and SIF are applied to a linear mechatronic system and the results are compared. In this case, a linear aerospace flight surface actuator, referred to as the electrohydrostatic actuator (EHA) is considered. The EHA is a type of mechatronic system that modifies the flight surface or flap of an aircraft to control aircraft flight motion. In the literature, the system has been modelled under two operating modes: normal conditions, and faulty conditions.

Similar to the state space equations defined in (2.1) and (2.2), the EHA system is defined respectively as follows:

$$x_{k+1} = \begin{bmatrix} 1 & T & 0 \\ 0 & 1 & T \\ -557 & -28.6 & 0.94 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ 557 \end{bmatrix} u_k + w_k \tag{4.1}$$

$$z_{k+1} = Cx_{k+1} + v_{k+1} \tag{4.2}$$

where the sample rate T is defined as $1\,ms$, C refers to the measurement matrix which is an identity matrix of dimension $m \times m$ or 3×3, and u is the controller input for the system (in this case, defined as Fig. 1) that drives the desired trajectory. The system and measurement noises (w and v) are normally distributed with zero mean and covariance's Q and R defined by (4.3) and (4.4), respectively.

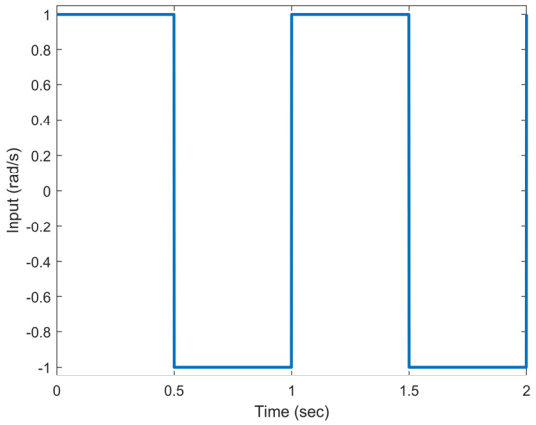


Figure 2. Control signal (rad/s) used as the input for the linear EHA simulation described by (5.1) to (5.4).

$$Q = diag([10^{-5} \quad 10^{-3} \quad 0.1]) \tag{4.3}$$

$$R = diag([10^{-4} \quad 10^{-2} \quad 1]) \tag{4.4}$$

The three states defined by the system (4.1) are kinematic EHA states, or position, velocity, and acceleration. Note that in this simulation, the states, measurements, and estimates were all set to zero. The initial state error covariance was set to $P_{0|0} = 10Q$. Furthermore, the boundary layers were tuned manually to yield the smallest estimation error, and were set to

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

$\psi = \delta = [0.05 \quad 1 \quad 0.5]^T$. The simulation was coded in the MATLAB environment.

The results of applying the KF, SVSF, and SIF strategies on the linear EHA are shown in Fig. 3. This plot shows the EHA position with time with the overlapping estimates. All of the filters were able to provide good estimates of the states. However, as expected, the KF provided the best results as it yields the optimal solution when the system and measurements are known. One measure of performance used in this study is the root mean squared error (RMSE) which is defined by (4.5) where n is the number of time steps. The results are summarized in Tab. I.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{n}} \tag{4.5}$$

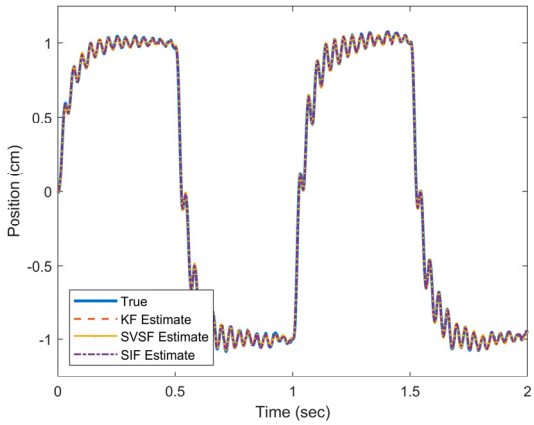


Figure 3. True and estimated position values for the linear EHA system example under normal operating conditions. The results are nearly identical so appear overlapping. However, note that for this case, the KF yields the best results which is expected since it is an optimal filter for known, linear systems and measurements.

TABLE I
EHA RMSE RESULTS: NORMAL CASE

State	KF	SVSF	SIF
Position (m)	3.83×10^{-3}	6.29×10^{-3}	5.92×10^{-3}
Vel. (m/s)	5.11×10^{-2}	6.38×10^{-2}	5.75×10^{-2}
Accel. (m/s^2)	9.00×10^{-1}	9.71×10^{-1}	9.62×10^{-1}

For completeness, the true and estimated state values for the velocity and acceleration are shown in Figs. 4 and 5.

To demonstrate the robustness of the variable structure and sliding mode strategies, consider the case when the system has a fault injected half-way through the simulation (at $t = 1$ sec). In this faulty case, the linear system state equation used by the system is changed as follows:

$$x_{k+1} = \begin{bmatrix} 1 & T & 0 \\ 0 & 1 & T \\ -240 & -28 & 0.94 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ 557 \end{bmatrix} u_k + w_k \tag{4.6}$$

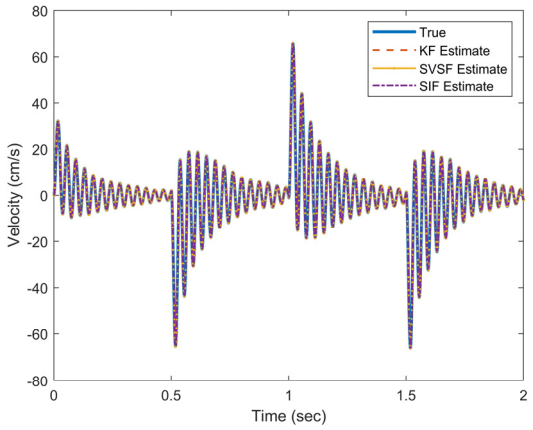


Figure 4. True and estimated velocity values for the linear EHA system example under normal operating conditions. The results are nearly identical so appear overlapping.

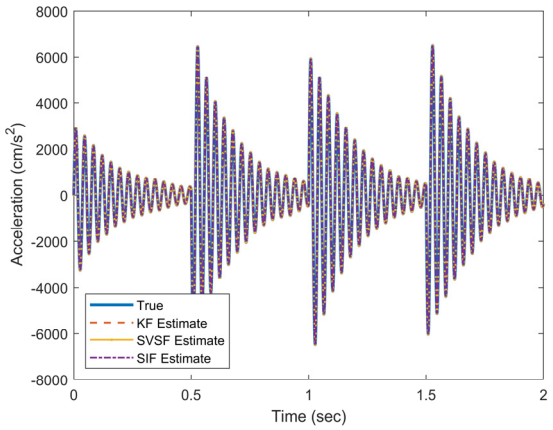


Figure 5. True and estimated acceleration values for the linear EHA system example under normal operating conditions. The results are nearly identical so appear overlapping.

Note that the filters do not use (4.6) but still use (4.1) to define the system dynamics. This causes model-mismatch or modeling uncertainties to occur. The results of the modeling uncertainty and its effects on estimating the EHA position are shown in Fig. 6. At one second, the fault is injection into the system. The model mismatch causes the KF to deviate from the true state trajectory, yielding poor estimates of the true position. However, both the SVSF and SIF were able to perform relatively well, and were bounded to the true state trajectory. This was due to the inherent stability caused by the switching effects and the defined boundary layers within their respective gains.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

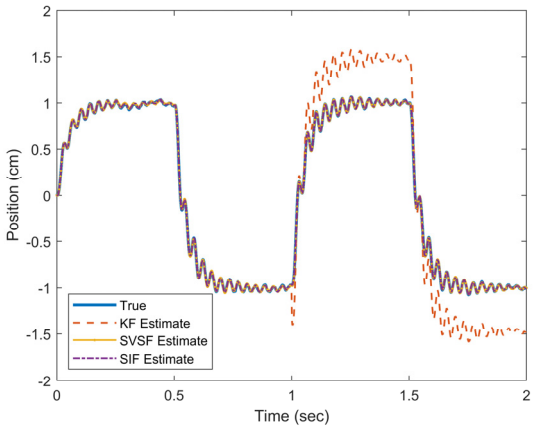


Figure 6. True and estimated position values for the linear EHA system example under faulty operating conditions. The negative effects of modeling uncertainity on the KF is clearly apparent after one second. The SVSF and SIF were robust to the model mismach and yielded relatively good estimates.

TABLE II
EHA RMSE RESULTS: FAULTY CASE

State	KF	SVSF	SIF
Position (m)	3.06×10^{-1}	6.42×10^{-3}	6.03×10^{-3}
Vel. (m/s)	3.44	6.67×10^{-2}	5.89×10^{-2}
Accel. (m/s^2)	17.69	9.98×10^{-1}	9.97×10^{-1}

The RMSE results for the faulty case are shown in Tab. II. Under the presence of a fault, the SVSF and SIF perform slightly worse than the normal case, which is to be expected. However, the KF was unable to recover from the modeling uncertainty and yielded significantly worse performance which would have a significant impact on aircraft flight performance. This was expected as the KF is derived based on the assumption that the system is known, and when that assumption is violated the KF yields poor and sometimes unstable results. For completeness, the true and estimated state values for the velocity and acceleration under faulty conditions are shown in Figs. 7 and 8.

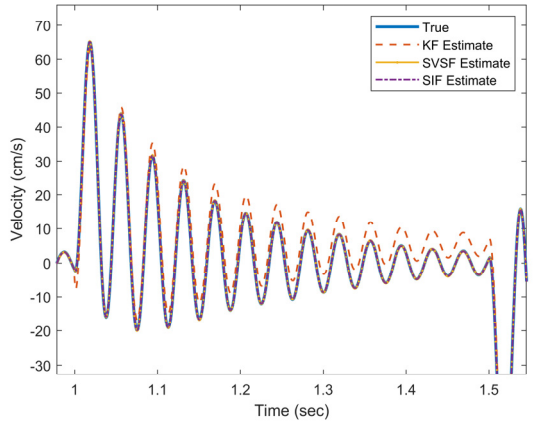


Figure 7. True and estimated velocity values for the linear EHA system example under faulty operating conditions. Note that the shown results are between about 1 and 1.5 seconds to illustrate the KF deviations.

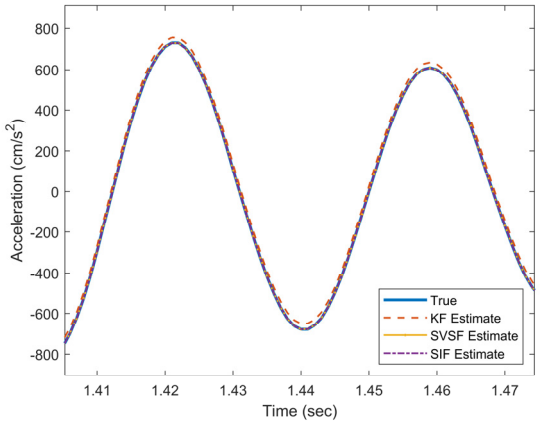


Figure 8. True and estimated acceleration values for the linear EHA system example under faulty operating conditions. Note that the results were zoomed-in to illustrate the KF deviations.

Figure 9 further illustrates the presence of the modeling uncertainty injected at 1 second in the simulation. The KF was unable to recover from the system change, whereas the SVSF and SIF were still able to provide a good estimate. This further demonstrates the robustness of the SVSF and SIF for modeling uncertainties.

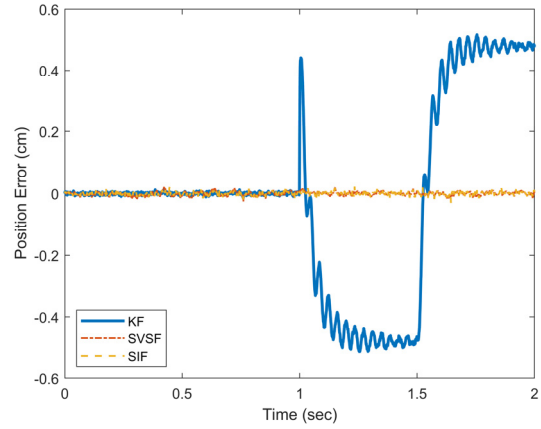


Figure 9. Position estimation error for the KF, SVSF, and SIF as shown in the faulty case (i.e., modeling uncertainty).

V. CONCLUSIONS

In this paper, a study of estimation strategies based on variable structure and sliding mode theory was performed. The smooth variable structure filter (SVSF) and the new sliding innovation filter (SIF) are based on similar sliding mode concepts but with some notable differences. The main difference is in the formulation of the gain and requirement of information from previous time steps. The SVSF and SIF gains are inherently robust to modeling uncertainties and disturbances due to the switching effects of the saturation function. A linear mechatronic system was studied where the SVSF and SIF were compared with the well-known Kalman filter (KF). The results indicate that although both the SVSF and SIF provide robust estimates to faults, the SIF formulation provides slightly more accurate estimates while maintaining robustness. Furthermore, the SIF estimation process is determined to be simpler and less computationally complex.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

REFERENCES

[1] H. H. Afshari, S. A. Gadsden and S. R. Habibi, "Gaussian Filters for Parameter and State Estimation: A General Review and Recent Trends," *Signal Processing,* vol. 135, pp. 218-238, 2017.

[2] S. A. Gadsden, "Smooth Variable Structure Filtering: Theory and Applications," McMaster University (PhD Thesis), Hamilton, Ontario, 2011.

[3] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE,* vol. 92, no. 3, 2004.

[4] M. Al-Shabi, "The General Toeplitz/Observability SVSF," Hamilton, Ontario, 2011.

[5] Z. Xu, S. X. Yang and S. A. Gadsden, "Enhanced Bioinspired Backstepping Control for a Mobile Robot with Unscented Kalman Filter," *IEEE Access,* vol. 8, pp. 125899-125908, 2020.

[6] J. E. Potter and R. G. Stern, "Statistical Filtering of Space Navigation Measurements," in *Proceedings of 1963 AIAA Guidance and Control Conference*, New York, 1963.

[7] N. A. Carlson, "Fast Triangular Formulation of the Square Root Filter," *AIAA Journal,* vol. 11, no. 9, pp. 1259-1265, 1973.

[8] G. J. Bierman, Factorization Methods for Discrete Sequential Estimation, New York: Academic Press, Inc., 1977.

[9] R. Van der Merwe and E. A. Wan, "The Square-Root Unscented Kalman Filter for State and Parameter-Estimation," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2001.

[10] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, Numerical Recipes: The Art of Scientific Computing, New York: Cambridge University Press, 2007.

[11] Y. Bar-Shalom, X. Rong Li and T. Kirubarajan, Estimation with Applications to Tracking and Navigation, New York: John Wiley and Sons, Inc., 2001.

[12] D. Simon, Optimal State Estimation: Kalman, H-Infinity, and Nonlinear Approaches, Wiley-Interscience, 2006.

[13] S. A. Gadsden and A. S. Lee, "Advances of the Smooth Variable Structure Filter: Square-Root and Two-Pass Formulations," *Journal of Applied Remote Sensing,* vol. 11, no. 1, 2017.

[14] S. R. Habibi, "The Smooth Variable Structure Filter," *Proceedings of the IEEE,* vol. 95, no. 5, pp. 1026-1059, 2007.

[15] S. A. Gadsden and S. R. Habibi, "Derivation of an Optimal Boundary Layer Width for the Smooth Variable Structure Filter," in *ASME/IEEE American Control Conference (ACC)*, San Francisco, California, 2011.

[16] S. A. Gadsden, Y. Song and S. R. Habibi, "Novel Model-Based Estimators for the Purposes of Fault Detection and Diagnosis," *IEEE/ASME Transactions on Mechatronics,* vol. 18, no. 4, pp. 1237-1249, 2013.

[17] M. Al-Shabi, S. A. Gadsden and S. R. Habibi, "Kalman Filtering Strategies Utilizing the Chattering Effects of the Smooth Variable Structure Filter," *Signal Processing,* vol. 93, no. 2, pp. 420-431, 2013.

[18] H. H. Afshari, S. A. Gadsden and S. R. Habibi, "Robust Fault Diagnosis of an Electro-Hydrostatic Actuator using the Novel Optimal Second-Order SVSF and IMM Strategy," *International Journal of Fluid Power,* vol. 15, no. 3, pp. 181-196, 2014.

[19] H. H. Afshari, "The 2nd-Order Smooth Variable Structure Filter (2nd-SVSF) for State Estimation: Theory and Applications," McMaster University (PhD Thesis), Hamilton, Ontario, 2014.

[20] H. H. Afshari, S. A. Gadsden and S. R. Habibi, "A Nonlinear Second-Order Filter for State Estimation of Uncertain Systems," *Signal Processing,* vol. 155, pp. 182-192, 2019.

[21] M. Attari, "SVSF Estimation for Target Tracking with Measurement Origin Uncertainty," McMaster University (PhD Thesis), Hamilton, Ontario, 2016.

[22] S. A. Gadsden, M. Al-Shabi, I. Arasaratnam and S. R. Habibi, "Combined Cubature Kalman and Smooth Variable Structure Filtering: A Robust Estimation Strategy," *Signal Processing,* vol. 96, no. B, pp. 290-299, 2014.

[23] M. Al-Shabi, K. S. Hatamleh, S. A. Gadsden, B. Soudan and A. El-Nady, "Robust Nonlinear Control and Estimation of an PRRR Robotic Arm," *International Journal of Robotics and Automation,* vol. 34, no. 6, 2019.

[24] J. Kim, K. Chang, B. Schwarz, A. S. Lee, S. A. Gadsden and M. Al-Shabi, "Dynamic model and motion control of a robotic manipulator," *Journal of Robotics, Networking and Artificial Life,* vol. 4, no. 2, pp. 138-141, 2017.

[25] S. A. Gadsden and S. R. Habibi, "A New Robust Filtering Strategy for Linear Systems," *ASME Journal of Dynamic Systems, Measurement, and Control,* vol. 135, no. 1, 2013.

[26] W. Youn and S. A. Gadsden, "Combined Quaternion-Based Error State Kalman Filtering and Smooth Variable Structure Filtering for Robust Attitude Estimation," *IEEE Access,* vol. 7, pp. 148989-149004, 2019.

[27] S. A. Gadsden, S. R. Habibi and T. Kirubarajan, "Kalman and Smooth Variable Structure Filters for Robust Estimation," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 50, no. 2, pp. 1038-1050, 2014.

[28] J. Goodman, S. A. Wilkerson, C. Eggleton and S. A. Gadsden, "A multiple model adaptive SVSF-KF estimation strategy," in *SPIE Signal Processing, Sensor/Information Fusion, and Target Recognition XXVIII*, Baltimore, Maryland, 2019.

[29] K. S. Hatamleh, M. Al-Shabi, A. Al-Ghasem and A. A. Asad, "Unmanned aerial vehicles parameter estimation using artificial neural networks and iterative bi-section shooting method," *Applied Soft Computing,* vol. 36, pp. 457-467, 2015.

[30] S. A. Gadsden and M. Al-Shabi, "The Sliding Innovation Filter," *IEEE Access,* vol. 8, pp. 96129-96138, 2020.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Robot Action Planning by Commonsense Knowledge in Human-Robot Collaborative Tasks

Christopher J. Conti
Deparment of Computer Science
Montclair State University
Montclair, NJ. USA
contic5@montclair.edu

Aparna S. Varde
Department of Computer Science
Montclair State University
Montclair, NJ, USA
vardea@montclair.edu

Weitian Wang
Department of Computer Science
Montclair State University
Montclair, NJ, USA
wangw@montclair.edu

Abstract*—Robotics and artificial intelligence (AI) span the broad realm of mechatronics in general. Humans and robots can collaborate with each other to enhance various tasks, e.g. in the vehicle industry. Facets of AI such as commonsense knowledge can play a significant role here. In this paper, we propose an approach for human-robot collaboration such that it leverages commonsense knowledge to develop models for the simulation of assembly tasks in real-world applications. We consider modeling based on relevant attributes, e.g. weight and stability of parts. The proposed approach thereby entails human-robot interaction, mathematical modeling, both semantics and pragmatics in commonsense knowledge, as well as challenges specific to the application domain. We describe our approach, focusing on the mathematical modeling, and conduct our experimental evaluation using simulation tasks. Experimental results indicate that the proposed approach yields better outcomes than humans or robots working alone, which is in line with other such claims in the field of human-robot collaboration. This work sets the stage for real robot applications based on the results of our simulation tasks.

***Keywords*—artificial intelligence; commonsense knowledge; robotics; human-robot interaction; mathematical modeling; simulation studies**

I. INTRODUCTION

Human-robot collaboration (HRC), where robots work with humans on tasks, is an important aspect of various industries such as manufacturing. Rather than removing humans from manufacturing, collaboration involves robots supporting humans [1]. Traditional robots typically work in a separate space while collaborative robots can work in the same space as humans[2], [3]. Collaborative robots are also modifiable, which means that they can be used for disparate tasks, unlike traditional robots [4-6]. Fig. 1 illustrates a collaborative robot used in the context of vehicle assembly. For robots to be able to collaborate with human beings optimally, it is useful for them to have commonsense knowledge analogous to humans.

Commonsense Knowledge (CSK) involves a basic, intuitive understanding of real-world objects, their properties and associations, and the manner in which they interact with each other. Humans inherently use CSK while robots need to be programmed for using CSK in order to handle everyday real-world situations in a manner analogous to humans [7]. There are commonsense knowledge bases developed in the literature, including the classical ConcepNet, Cyc, DBpedia, WordNet etc. as surveyed in [7] and the more recently built WebChild [8]. Fig. 2 presents a screenshot of the WebChild commonsense browser. It provides explanation of everyday concepts with their description, type, domain, property, location etc. Here we see an example of the concept truck which is a type of "motor vehicle" having its typical activity as "drive" and location as "road". Such commonsense knowledge bases provide useful information that can be programmed into robots in order to make them smarter and more humanlike, especially in autonomous decision-making.

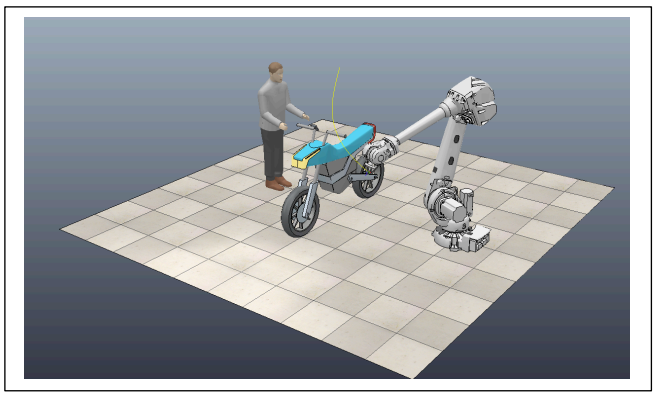


Fig 1. Example of human-robot collaboration in vehicle assembly

Fig 2. Example of the WebChild commonsense browser

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

This study is motivated by adding commonsense knowledge to robots so that they can handle human-robot collaboration more effectively and implement it in useful applications. Given this background and motivation, we define the problem in our work and explain our proposed approach with experimentation and discussion.

Problem definition: The problem addressed in this paper is defined with the following two goals.
1. Develop an approach for human-robot collaboration that incorporates commonsense knowledge;
2. Deploy the approach in the context of robot action planning for object assembly to demonstrate its effectiveness.
We propose our solution approach accordingly and deploy it with simulations for human-robot collaborative tasks.

In order to describe our proposed approach, we first present an overview of the related literature in the areas of both commonsense knowledge and human-robot collaboration. This brief literature survey helps to gain an understanding of our approach with respect to the state-of-the-art. Some studies might be inadvertently touching upon commonsense capabilities within their experimentation of humans and robots working together. Some works in the literature outline that commonsense semantics is important and useful in the realms of AI and robotics. Yet, to the best of our knowledge, our paper is among the first works to explicitly harness commonsense knowledge in human-robot collaboration by proposing an approach for CSK-based-HRC, and deploying it in robot action planning for object assembly. Challenges in this work include capturing subtle nuances of commonsense that may be trivial for humans but are difficult for robots; and conducting adequate collaborative tasks such that satisfactory results are achieved in targeted applications.

The rest of this paper is organized as follows. Section II overviews related work in the area. Section III describes our proposed approach of commonsense knowledge based human-robot collaboration, encompassing the corresponding knowledge base development and mathematical modeling. Section IV provides a summary of our experimental evaluation for human-robot collaborative tasks using simulation, along with suitable discussion in an application-specific context. Section V states the conclusions of this paper with perspectives, and lists future work in the area.

II. Related Work

Commonsense knowledge focuses on day-to-day entities, how they are related and how they interact [7]. It includes knowledge that is obvious to humans, e.g. they know the fact that the sky is blue (even though that is not connected to their work specialty). It also includes how attributes are connected, such as the fact that gray skies indicate that rain is probable. Machines are very competent at memorizing a plethora of facts and information, but have a much more difficult time with commonsense knowledge, especially in the absence of explicit prior training [8].

Robots need commonsense knowledge to operate in real-world spaces. Automated vehicles for instance, need to be able process what they detect in order to handle dynamic driving environments. For example, in 2016, a Tesla vehicle collided with a truck in motion after mistaking it for an overpass, leading to a fatal accident [9]. A good human driver would easily be able to distinguish between a truck and an overpass, e.g. an overpass is stationary while a truck can move, and this fact is known to humans intuitively due to common sense [10]. Such facts can be programmed via CSK into vehicles for automated driving using information from commonsense knowledge bases (see Fig. 2). Additionally, a vehicle can use such CSK to track objects even after they leave its view since objects do not suddenly disappear from the road, as is obvious to humans [11]. Aspects such as these motivate the need for automated vehicles to be embedded with CSK [10] so that they can be closer to the thresholds of human cognition and conduct better decision-making accordingly.

Human-robot collaboration is quite important and beneficial in Artificial Intelligence (AI) [1], [12], [13]. *Collaborative robots* are defined as robots that can work alongside human beings within the same space rather than needing their own separate space [2], [3]. This cooperation combines the best elements of humans and robots, such as the strength and repetition skills of robots, along with the judgement and adaptation skills of humans [2]. Traditional industrial robots require extra guarding and equipment for safety, thereby increasing their cost and bulkiness while decreasing their flexibility. Additionally, traditional robots cannot be modified while collaborative robots can be modified to handle different tasks due to their flexibility.

Robots can be used for manufacturing tasks in order to increase safety and productivity in the workplace [6]. Robots are typically used for tasks such as welding, assembly and paint spraying. Robot welding is useful since it produces superior products while making the work safer and less monotonous for humans. Likewise, robot assembly is useful since the costs are lowered while throughput and consistency are increased. Using robots can benefit humans while improving the bottom margin. Robots using paint spray not only improves consistency as well as throughput, but also keeps human operators away from hazardous environments. The use of industrial robots therefore results in more safe work environments and thus makes positive impacts on production as a whole.

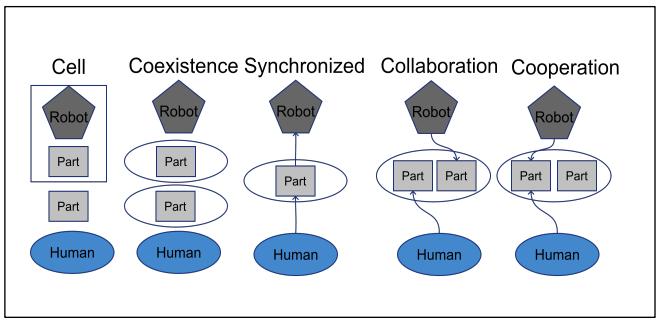


Fig. 3. Types of human-robot collaboration

As illustrated in Fig. 3, where each dark gray pentagon depicts a “Robot”, each white circle indicates a workspace, each light gray square represents a “Part” and each blue circle portrays a “Human”; there are four main types of human-robot collaboration. These are coexistence, synchronized, cooperation, and (full) collaboration [2]. For *coexistence,* humans and robots work in the same space without interacting with each other. In the *synchronized* approach,

humans and robots work in the same space, but at different times. For *cooperation*, humans and robots work together in the same space simultaneously, but work on different tasks. For the full *collaboration* approach, often referred to as just "collaboration", humans and robots actually work together on the same tasks, with their actions affecting each other [2]. These different types of collaboration in the realm of HRC can be useful for different situations, since humans may perform better in some kinds of tasks while robots may work better in others. Human-robot collaboration thus has various forms and can be useful for many companies, especially in manufacturing applications.

Another example of useful human-robot collaboration is a robot helping out with tree fruit farming [14]. The robot assists human beings at different tasks by utilizing its three modes, Mule Mode, Pace Mode and Scaffold Mode. The Mule Mode assists with tasks such as harvesting while following the workers. The Pace Mode performs a specific task over a designated area. The Scaffold Mode has the robot become a scalpel that the humans can stand on, in order to execute tasks. It has been observed in real experiments with these different modes that workers standing on the robot in the Scaffold Mode have been able to trim fruit trees twice as fast as workers using ladders [14]. The results demonstrate that humans when cooperating with robots can be much more productive than when working by themselves.

Considering this overview of related work on human-robot collaboration entailing its techniques and applications, we present our proposed approach in this paper. This is on commonsense knowledge based human-robot collaboration and is described next.

III. Proposed Approach: CSK-Based-HRC

In this paper, we propose an approach that leverages the fundamental aspects of commonsense knowledge (CSK) to conduct human-robot collaboration (HRC) with specific deployment for robot action planning in object assembly. This approach is described along with mathematical modeling using attributes such as size, weight and distance in terms of parts in object assembly. It is explained in three subsections: framework of the approach, development of a knowledge base; and mathematical modeling, respectively.

A. Framework of the Approach

The two main types of commonsense knowledge are pragmatics, which relates to general world knowledge, and semantics, which relates to the specific context. An example of useful pragmatic knowledge is a delivery robot moving slower when in a dark area since people will have more trouble seeing it. An example of useful semantic knowledge is a driving robot adjusting its movements during heavy precipitation to avoid hydroplaning.

Throughout this paper, the term *"parts"* will be used to refer to individual components that make up a greater whole, while the term *"object"* is used to refer to the final result of combining parts in a pre-determined manner. Based on this, we explain object assembly using HRC.

The human-robot collaboration approach for assembling objects in this work uses commonsense knowledge so that the robots can more effectively select and move parts. The robot arm should prioritize heavier parts since humans prefer to carry lighter parts because that is easier. At the same time, the robot arm should not prioritize carrying fragile/unstable parts since the human may not trust it to handle those parts correctly. The robot should carry parts that are made of glass so humans do not cut themselves. These assembly tasks are often being repeated throughout the day, hence it is useful to make them easier for humans. In this collaborative approach, we deploy the following **fundamental premises in CSK**.

1. *Humans prefer to carry lighter parts due to ease*
2. *Robots may not be trusted with unstable parts*
3. *Glass parts can possibly be dangerous to humans*

Note that premises 2 and 3 might pose conflicts, since a glass part can be unstable. Hence, mathematical modeling of the corresponding attributes plays a role in decision-making, i.e. whether the stability of the part is more significant or its danger based on their respective values and priorities.

Given such premises, our proposed approach for human-robot collaboration is illustrated in Fig. 4. The human and robot execution affect the workspace that is used to determine the robot's next action. When determining the overall cost for parts, with reference to this figure, the actors' (human / robot) actions are considered. The result of the cost function, based on CSK-relevant attributes such as stability and weight, is used to determine robot execution.

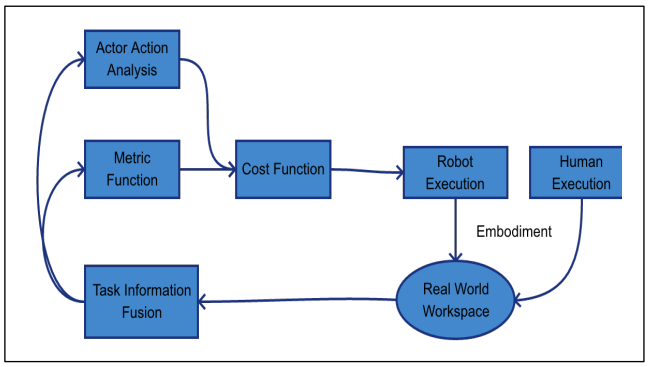


Fig. 4. Framework of proposed approach in object assembly

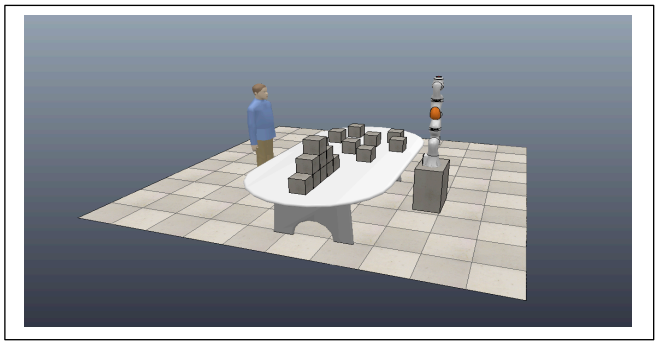


Fig. 5. Human and robot cooperate to combine parts to form the pre-determined object

The HRC system described herewith based on these CSK premises is designed to handle building a vehicle, with the initial parts being nearby and combined to form the final object, similar to a manner shown in Fig. 5. This figure presents a relevant excerpt of humans and robots collaborating with each other in order to build the concerned object.

B. Development of a Knowledge Base

The knowledge base (KB) developed in this paper focuses on human and robot priorities for certain attributes, such as weight, stability, danger and distance (relative to part's final position). Humans want to minimize certain attributes such as weight and distance. Due to that, robots should maximize those attributes. Humans may not necessarily trust robots, so humans will take parts that are less stable to avoid breaking them, thus robots should prioritize parts that are more stable.

Hence, the commonsense priorities embedded in our proposed CSK-based-HRC approach are as follows. These approximately correspond to the respective fundamental premises in CSK outlined in the previous subsection.

1. *Distance refers to distance traveled by the part, weight refers to heaviness of the part carried and size refers to the dimensions of a part. Humans prefer to minimize the values of distance, weight and size of parts in object assembly, thus robots should preferably maximize these values in the respective assembly tasks.*
2. *Stability refers to how stable or durable a part is, e.g. a clay piece would have lower stability than a wooden piece since the clay piece would break more easily. Humans prefer to handle less stable parts since robots cannot always be trusted with these.*
3. *Danger refers to how potentially dangerous or risky the parts are, e.g. scissors would have a greater danger value than sheets of paper, due to their sharpness. Humans prefer to avoid dangerous parts while robots can handle these since they do not pose any risk to robots.*

Accordingly, the CSK-based attributes (corresponding to these priorities) and their ranges of values are depicted in TABLE I herewith, as coded in the KB. The columns here indicate the CSK attributes. The first row defines the minimal human attribute values while the second row defines its maximal ones; the third row defines the maximal robot attribute values, and the fourth row defines its minimal ones.

TABLE I: RANGES OF CSK-BASED ATTRIBUTE VALUES FOR HUMANS AND ROBOTS

	Distance (cm)	Weight (kg)	Danger (level)	Stability (level)
Min for H	5	1	0	0
Max for H	200	50	30	30
Max for R	200	50	30	30
Min for R	5	1	0	0

The robot's maximization priorities are designed to mirror the human's minimization priorities and vice-versa. This allows humans to work with parts they prefer and are better at working with, resulting in object assembly being faster, safer and more effective. Robots can then handle those parts that humans have more difficulty moving, such as heavy, large or fragile parts. The priorities are weighted, such that danger is one and a half times as important as any of the other attributes. The other attributes have equal priorities under this commonsense priority system. This means that a part may be selected by the robot if its danger is higher and it is extremely close versus a part that is not dangerous, but is far away. This is because it is important to avoid danger to humans in HRC tasks. In other words, safety is the top priority. More on this will be included in the next subsection.

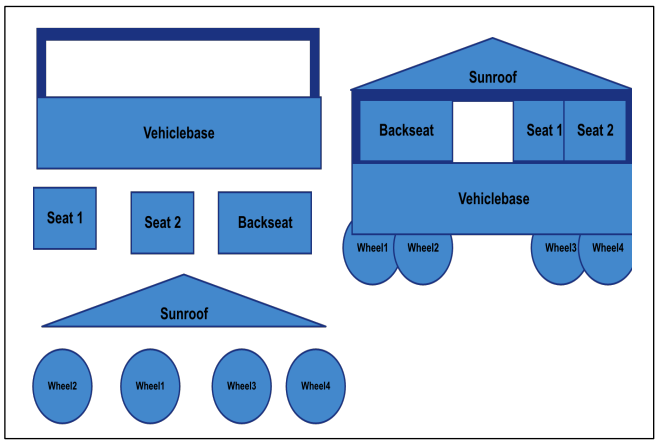


Fig. 6. Car parts in initial state (left) and final state (right) in vehicle assembly

Fig. 6 depicts an example of car parts in their initial and final state during vehicle assembly using HRC. This is achieved in the given domain of vehicle manufacturing based on a KB design analogous to that explained herewith, and the mathematical modeling as described next.

C. Mathematical Modeling

In our mathematical modeling and simulation, the arm movement is defined as follows. First, arms lock onto a part and indicate that they are conducting this locking. This prevents other arms from moving towards the same part. Arms then move to grab the part they locked onto and then move to place these where they connect with other parts. This process is repeated continuously until there are no remaining parts. The process of arm movement is formally described in *Algorithm 1* in the form of simplified code. The input to this algorithm includes the real-time parts along with their properties, positions and priorities. The output entails placing the parts in their correct positions before the robot action planning for the actual assembly.

Algorithm 1: Simplified code for robot arm movement

Input: The real-time parts and their properties, and the positions and priorities of arms
Output: The parts being placed in their correct positions

```
#arm.holdingpart starts as false
for arm in arms:
    if not(arm.lockedon):
        bestpart[arm] =arm.determinebestpart()
        bestconnect[arm]= arm.determinebestconnect()
        arm.lockonto(bestpart)
    if(arm.lockedon):
        if(not arm.holdingpart):
            arm.movetopart(bestpart)
            arm.pickup(bestpart)
        else:
            arm.movetoconnect( bestconnect[arm])
        arm.placepart(bestpart)
```

For robot action planning, consider that humans and robots collaborate in the same workspace with the same set of parts denoted as *P,* and commonsense priorities denoted as *C.* Thus

each *P* refers to a part such as wheel1, seat1 etc. while each *C* refers to a CSK priority such as distance, weight etc. Note that danger gets 1.5 times higher priority than the mean of all the other attributes. Here, *C(danger)* represents the CSK priority of this attribute while *C(x) for x = 1 to m* represents the priorities of each of the other attributes. We have 4 attributes here, but there could be more (based on the application). Furthermore, P_i represents the i^{th} part in the set *P* while $S(p_i)$ represents the overall score for moving a part to its correct position. This score is calculated by comparing the attributes of a part against the maximum and minimum values of the attributes of all the parts, along with the CSK priorities. P_s represents the selected part while *t* represents the total number of parts. Thus, human and robot arms select their next part to move using Equations (1) to (3).

$$C(danger) = \frac{1.5}{m}\sum_{x=1}^{m} C(x) \quad -(1)$$

$$S = \{S(p_i) \mid 0 \le i \le t-1\} \quad -(2)$$

$$P_s = P\big(\text{argmax}(S)\big) \quad -(3)$$

An algorithm is used in order to calculate which part to select, i.e. this constitutes robot action planning based on a cost function that embodies these commonsense priorities. Arms have minimization priority and there is maximization priority for each attribute analyzed. This algorithm is presented here as *Algorithm 2*. The input to this algorithm is the same as that for *Algorithm 1* while its output is: the part that the current arm will target next in robot action planning.

Algorithm 2: Robot action planning based on cost function

Input: The real-time parts and their properties, and the positions and priorities of arms
Output: The part that the current arm will target next

```
1. maxscore=0
2. selectedpart=None
3. for part in parts:
4.                 for pri in priorities:
5.                         curattrval=part.pri
6.                         if(curattrval >maxattrvals [pri]):
7.                                 maxattrvals [pri]= curattrval
8.                         if(curattrval < minattrvals [pri]):
9.                                 minattrvals [pri]= curattrval
10.        for part in parts:
11.                partscore=0
12.                for pri in priorities:
13.                        partscore+=
prioritiesmax[currentarm][pri]* curattrval / maxattrvals [pri]
14.                        partscore+=
prioritiesmin[currentarm][pri] * minattrvals [pri] / curattrval
15.                if(partscore>maxscore):
16.                        maxscore=partscore
17.                        selectedpart=part
18.        return selectedpart
```

These algorithms are used for the execution of tasks in our proposed approach for human-robot collaboration. We now describe its experimental evaluation based on simulations.

IV. Experimental Evaluation

We summarize the evaluation of our CSK-based-HRC approach for object assembly herewith. In this paper, our focus is on simulation tasks. The results obtained from this work would be useful in conducting further experiments in laboratories using real robots in specific applications. We explain our simulation tasks along with the experimental results followed by a discussion on the evaluation.

A. Simulation Tasks

The given task here is to assemble a vehicle with the following parts: wheel 1, wheel 2, wheel 3, wheel 4, vehicle-base, seat 1, seat 2, backseat and sunroof. The parts are then combined in order to form a vehicle (see Fig. 6). When making comparisons, all the parts maintain the same initial positions, though the system can give the parts random initial positions. A human and a robot are both working on this task using one arm that can hold one part at a time and are given priorities for selecting parts.

The system has been tested with **four different priority options** as follows.

- *Option 1: CSK priorities*
- *Option 2: No priorities*
- *Option 3: Human only using priorities (with CSK) and robot not using priorities*
- *Option 4: Closest part priorities (with CSK)*

The main goal here is to make work easier for humans while allowing the overall assembly time to remain fairly consistent. For synchronous execution, both arms move and place parts at the same time. For asynchronous execution, only one arm moves and places a part at a time, with the arms switching every time a part is placed. In our experiments, distance is measured in cm, weight in kg, and time in seconds. Stability and danger do not have specific units, we measure them by scaled levels. We tabulate our experimental results next, presenting the highlights of the executions.

TABLE II: VALUES OF HUMAN ATTRIBUTES AND THE TOTAL CO-ASSEMBLY TIME FOR ASYNCHRONOUS EXECUTION

Attributes	Distance	Weight	Stability	Danger	Time
CSK priorities	100	30	12	10	53
No priorities	127	30	14	14	55
Human only p.	116	30	12	22	53
Closest part p.	91	45	14	26	49

TABLE III: VALUES OF HUMAN ATTRIBUTES AND THE TOTAL CO-ASSEMBLY TIME FOR SYNCHRONOUS EXECUTION

Attributes	Distance	Weight	Stability	Danger	Time
CSK priorities	84	35	10	11	29
No priorities	104	25	8	20	32
Human only p.	108	35	10	23	33
Closest part p.	91	45	14	26	27

TABLES II and III present the observed values of human attributes and the total co-assembly time (for humans and robots combined), measured in seconds for synchronous and asynchronous execution respectively. In both these tables, the rows “CSK priorities” … “Closest part p.” respectively refer to the four different priority options used for testing as described here. The columns are explained as follows.

"Distance" refers to how far the human arm traveled in cm and "Weight" refers to the heaviness of the part in kg that the human arm carried. "Stability" and "Danger" refer to the total scaled levels of durability and risk respectively for the parts the human arm carried. (This is as per the explanation of commonsense priorities earlier). "Time" refers to the total duration in seconds to complete the entire task of co-assembly for both the human arm and robot arm combined.

These tables show the result of executing the human-robot collaboration model with a consistent set of data. It is observed that the results are quite similar for synchronous and asynchronous execution. The human commonsense priorities focus on minimizing travel distance, weight and danger, with danger minimization being the most prioritized. Due to this, it is observed that the human carries the least dangerous parts when using the CSK priorities. Human commonsense priorities focus on carrying the least stable parts because humans are more precise than robots. Note that minimizing danger and distance are prioritized over minimizing stability. It is observed that the task is completed faster with human and robot both using CSK priorities, as compared to the situation where the robot has no priorities or neither entity (human / robot) has priorities. It is also observed that the task is completed when using CSK priorities (Option 1) nearly as fast as when grabbing of the closest parts is prioritized (Option 4). The algorithm decides which parts to pick based on all of the attribute priorities, meaning that if a human has to choose between a very dangerous part that is close and a very safe part that is further away, the human arm will choose the safer part. The scoring system allows for prioritizing more than one attribute. The minimum distance is achieved with the closest part priority (Option 4) since minimizing distance is the only goal that is prioritized there. However, the weight and danger carried by humans in that case are both greater than for CSK priorities. In general, the CSK priorities are more effective, even if they may be slightly slower.

B. *Discussion on Evaluation*

Based on these simulation results, we find that using commonsense knowledge makes work easier for humans. In addition, humans are safer since they are less frequently carrying parts that are dangerous.

Conversely, in situations where things could be completed faster, more danger would be added; and when performing repeated tasks that can increase the chance of injury. While two humans could be used for assembling the parts into the final object, they would become more tired. Such assembly tasks are being repeated multiple times throughout a day in a real scenario, making this tiredness even more significant.

This system as of now assumes that heavier and less stable parts can be moved at the same speed as easier to move parts. However as per CSK, heavier parts would be moved slower that lighter ones; likewise less stable parts would also be moved slower than more stable ones. This means that in the future, adding such aspects of commonsense knowledge can be even more effective than shown in the current simulation.

Research in related areas also supports the results of our simulation experiments. Trust is an implicit commonsense factor that needs to be considered. Trust is essential for seamless human-robot collaboration and user adoption of autonomous systems, such as robot assistants [15]. If users overly trust a robot, there may be issues when it is given too many tasks to handle. If users insufficiently trust a robot, its productivity and usefulness will be reduced. Human trust in robots is a feature that can be managed so that the robots are most effective.

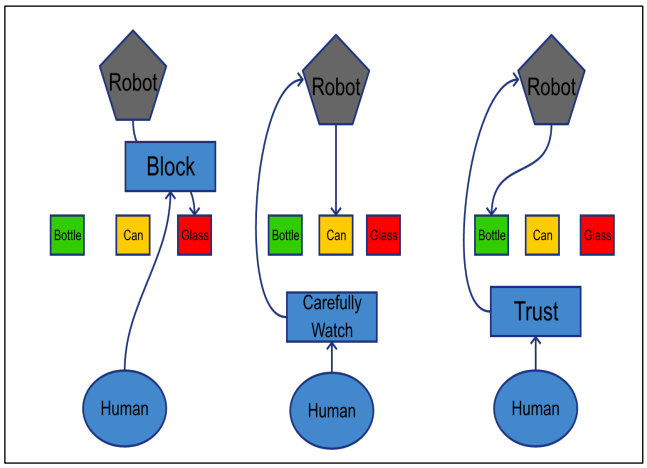


Fig. 7. A robot removing objects from a table with a human collaborating

In a study within this area, researchers had a human and a robot cooperate to clean off a table, with bottles, cans and wine glasses placed on it [15]. This is illustrated in Fig. 7 herewith. Robots that focused primarily on maximizing the reward removed the more fragile wine glass first, causing more human intervention. Robots that focused on trust removed bottles before attempting to remove the wine glass, which caused the humans to be less likely to intervene. Doing so maximized the reward, thereby showing how commonsense knowledge is inadvertently useful in these tasks.

Our experimental results corroborate such claims, hence making our study orthogonal to the literature. Moreover, our approach explicitly harnesses commonsense knowledge in its functioning with the appropriate KB development and mathematical modeling. Heretofore, CSK might have been implicitly noticed from the results of some studies while our work precisely deploys it in the proposed approach for CSK-based-HRC, specifically in the context of robot action planning in object assembly. This paper thus makes very interesting contributions to the field of human-robot collaboration and the more general realm of robotics.

V. Conclusions and Roadmap

In this paper, we address the issue of human-robot collaboration and present an approach for this based on commonsense knowledge, particularly with reference to robot action planning for object assembly. Our proposed approach for CSK-based-HRC is described here including its general framework, KB development and mathematical modeling. We present algorithms for robot arm movement and for robot action planning. We conduct simulation experiments for CSK-based-HRC in the application of vehicle assembly, useful in the automobile manufacturing industry. Our results prove that robots and humans working together guided by commonsense knowledge can be very effective, especially taking into account human safety and comfort. Our simulation tasks set

the stage for conducting real laboratory experiments in targeted applications. These are on our roadmap.

Currently, in this work the CSK priorities are kept static throughout the execution. Since human trust in robots can change as human-robot collaboration continues, designing modifiable priorities is valuable and can be considered in enhancements of this work. Future work would involve the robot interacting with a human during the simulation, as currently the human actions are simulated based on expected priorities. Trust is a factor that can be considered when modifying this simulation. Priority values for attributes such as stability can change as successful execution continues, which may allow robots to grab closer parts that are less stable. Additionally, the human priorities here are currently based on the developers' opinions. They can be improved by surveying targeted users and conducting more experiments. Through modification based on real-world data, this human-robot collaboration system can thus be enhanced.

Human-robot collaboration will play a key role in the future. Collaborative robots have features that allow them to be very useful. Rather than supplanting humans, robots can supplement humans in their work [1]. While humans can utilize careful judgement, can adapt and can depict creativity, robots can repeat tasks with more strength than humans, and can use their machine intelligence with their high memory and execution power [2]. It is notable to mention the suggestion of the famous chess grandmaster Garry Kasparov, mentioned at AAAI-2020 (international conference of the Association for Advancement of Artificial Intelligence), as well as stated in his book on Deep Thinking [16]. While Kasparov in 1997 as the world's then reigning chess champion lost a historic match to IBM's Deep Blue, today he prefers to have humans and robots work *together* rather than against each other, and head towards "Responsible Robotics" that combines "Machine Intelligence" and "Human Creativity". Kasparov states that "Humanity can rise to new heights with the *help* of our most extraordinary creations (robots) rather than fear them". This implies endorsement for human-robot collaboration (instead of just competition) which would be beneficial in many facets of life and useful in several real-world applications.

Our paper provides the two cents in this collaborative robotics area via our proposed CSK-based-HRC approach. This work makes contributions to AI and Robotics, thus bridging the area of Mechatronics as a whole. *The novelty of this paper, to the best of our knowledge, is that it is among the first works to focus on explicitly harnessing commonsense knowledge for human-robot collaboration, specifically in the context of robot action planning for object assembly.* It would be interesting to computer scientists and engineers, especially in AI and robotics, as well as domain experts in fields such as the smart manufacturing and intelligent transportation.

ACKNOWLEDGMENT

This work entails financial support from Montclair State University. Christopher J. Conti is funded by a graduate assistantship from the Department of Computer Science. Aparna S. Varde is supported by a doctoral faculty program from the Graduate School and a faculty scholarship program from Computer Science. Weitian Wang has startup funds from the College of Science and Mathematics.

REFERENCES

[1] K.-D. Thoben, S. Wiesner, and T. Wuest, "'Industrie 4.0' and smart manufacturing-a review of research issues and application examples," *International Journal of Automation Technology*, 2017, vol. 11, no. 1, pp. 4–16.

[2] J. Suchan and M. Bhatt, "Commonsense Scene Semantics for Cognitive Robotics: Towards Grounding Embodied Visuo-Locomotive Interactions," 2017, doi: 10.1109/ICCVW.2017.93.

[3] Y. Chen, W. Wang, Z. Abdollahi, Z. Wang, J. Shulte, V. Krowi, and Y. Jia, "A Robotic Lift Assister: A Smart Companion for Heavy Payload Transport and Manipulation in Automotive Assembly," *IEEE Robotics and Automation Magazine*, 2018, vol. 25, no. 2, pp. 107–119.

[4] D. E. Whitney, C. A. Lozinski, and J. M. Rourke, "Industrial robot forward calibration method and results," Journal of Dynamic Systems, Measurement and Control, 1986, vol. 108, no. 1, pp. 1-8.

[5] T. Pettersen, J. Pretlove, C. Skourup, T. Engedal, and T. Lokstad, "Augmented reality for programming industrial robots," in *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality,* 2003, pp. 319–320.

[6] J. Iqbal, R. U. Islam, S. Z. Abbas, A. A. Khan, and S. A. Ajwad, "Automating industrial tasks through mechatronic systems–A review of robotics in industrial perspective," *Teh. Vjesn.*, vol. 23, no. 3, pp. 917–924, 2016.

[7] N. Tandon, A. S. Varde, and G. de Melo, "Commonsense Knowledge in Machine Intelligence," *ACM SIGMOD Record,* 2017, vol. 46, no. 4, pp. 49-52.

[8] N. Tandon, G. De Melo, and G. Weikum, "WebChild 2.0: Fine-grained commonsense knowledge distillation," *The Annual Meeting of the Association for Computational Linguistics (ACL) conference* 2017, pp. 115-120.

[9] P. Pandey, M. Puri, and A. Varde, "Object Detection with Neural Models, Deep Learning and Common Sense to Aid Smart Mobility" *IEEE International Conference on Tools with Artificial Intelligence (ICTAI)* 2018, pp. 859-863.

[10] P. Persaud, A. S. Varde, and S. Robila, "Enhancing autonomous vehicles with commonsense: Smart mobility in smart cities," *IEEE International Conference on Tools with Artificial Intelligence (ICTAI)* 2017, pp. 1008-1012.

[11] E. Davis and G. Marcus, "Commonsense Reasoning and Commonsense Knowledge in Artificial Intelligence." [Online] Available: http://cs.nyu.edu/faculty/davise/papers/WS.html.

[12] W. Wang, R. Li, Y. Chen, Z. M. Diekel, and Y. Jia, "Facilitating Human–Robot Collaborative Tasks by Teaching-Learning-Collaboration From Human Demonstrations," *IEEE Transactions on Automation Science and Engineering*, 2018, vol. 16, no. 2, pp. 640–653.

[13] E. Matheson, R. Minto, E. G. G. Zampieri, M. Faccio, and G. Rosati, "Human-robot collaboration in manufacturing applications: A Review," *Robotics Journal, MDPI,* 2019, vol. 6, no.4, pp. 100, doi: 10.3390/robotics8040100.

[14] J. P. Vasconez, G. A. Kantor, and F. A. Auat Cheein, "Human–robot interaction in agriculture: A survey and current challenges," *Biosystems Engineering Journal*, 2019, pp. 35-48.

[15] M. Chen, S. Nikolaidis, H. Soh, D. Hsu, and S. Srinivasa, "Planning with Trust for Human-Robot Collaboration," Proceedings of the ACM / IEEE International Conference on Human-Robot Interaction, 2018, pp. 307-315.

[16] G. Kasparov, and M. Greengard, *Deep Thinking: Where Machine Intelligence Ends and Human Creativity Begins,* Public Affairs Publishers, Perseus Books, LLC, USA, 2017.

R-MnasNet: Reduced MnasNet for Computer Vision

Prasham Shah
IoT Collaboratory IUPUI
Department of Electrical and Computer Engineering
Purdue School of Engineering and Technology
Indianapolis, USA
pashah@purdue.edu

Mohamed El-Sharkawy
IoT Collaboratory IUPUI
Department of Electrical and Computer Engineering
Purdue School of Engineering and Technology
Indianapolis, USA
melshark@iupui.edu

***Abstract*—In Deep Learning, Convolutional Neural Networks (CNNs) are widely used for Computer Vision applications. With the advent of new technology, there is an inevitable necessity for CNNs to be computationally less expensive. It has become a key factor in determining its competence. CNN models must be compact in size and work efficiently when deployed on embedded systems. In order to achieve this goal, researchers have invented new algorithms which make CNNs lightweight yet accurate enough to be used for applications like object detection. In this paper, we have tried to do the same by modifying an architecture to make it compact with a fair trade-off between model size and accuracy. A new architecture, R-MnasNet (Reduced MnasNet), has been introduced which has a model size of 3 MB. It is trained on CIFAR-10 [4] and has a validation accuracy of 91.13%. Whereas the baseline architecture, MnasNet [1], has a model size of 12.7 MB with a validation accuracy of 80.8% when trained with CIFAR-10 dataset. R-MnasNet can be used on resource-constrained devices. It can be deployed on embedded systems for vision applications.**

***Index Terms*—Convolutional Neural Networks (CNNs), Computer Vision, R-MnasNet, CIFAR-10, MnasNet.**

I. Introduction

Convolutional neural networks (CNNs) have made noteworthy progress in computer vision applications like image classification [19], object detection [21] etc. Deep Neural Networks become slow because of large number of computations. It difficult t o d eploy s uch s tate-of-the-art C NN m odels on resource-constrained platforms.

As there are limited resources on mobile devices, researchers have been working on designing compact models. By reducing the depth of the network and utilizing less expensive convolutions, the models become more compact. Designing such models is very challenging as there must be a fair trade-off between model size and accuracy.

In this paper, a new architecture, R-MnasNet, is introduced which is designed to work efficiently w hen d eployed on such resource constrained platforms like mobile or embedded devices. This architecture has been derived from its baseline architecture MnasNet. Some modifications w ere m ade t o this baseline architecture to make it compact with a fair trade-off between model size and accuracy.

The baseline architecture is discussed in section 2. Section 3 introduces the new architecture and discusses the modifications made to the baseline architecture. Section 4 will acknowledge the hardware and software requirements. Section 5 shows the experimental results achieved while training the networks with its specifications and comparison between the models. Section 6 concludes the paper by discussing the future scope and utility of R-MnasNet.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

II. Prior Work

The efficiency of CNNs can be improved by designing hardware aware networks which have less computational cost and more accuracy. These models have performance like state-

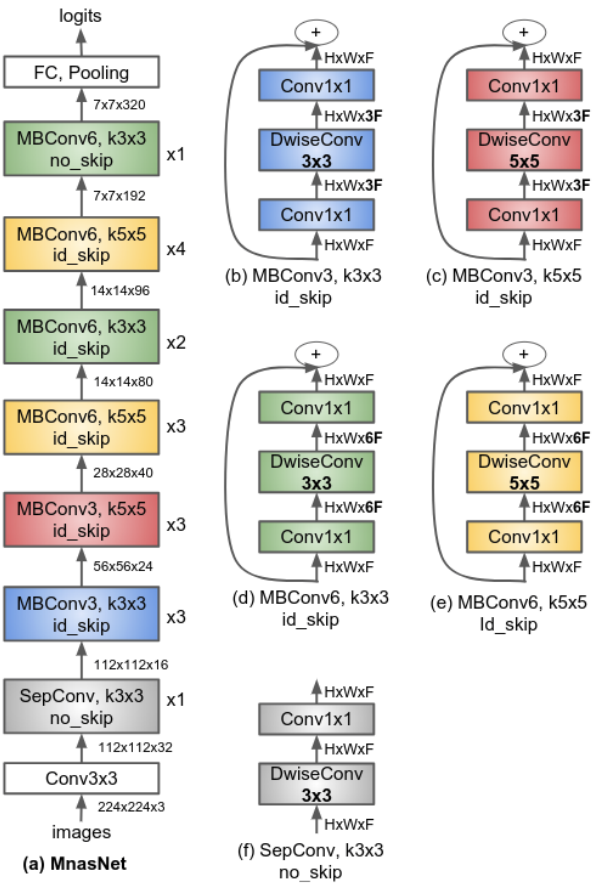


Fig. 1. MnasNet architecture

of-the-art CNN models. They are extensively used for mobile vision applications. MnasNet is such architecture which was designed to accomplish such goals.

A. MnasNet Architecture

They have introduced a neural architecture search approach which optimized accuracy and latency on mobile devices using reinforcement learning. By using this approach, they propose architectures based on real world latency and accuracy trade-offs. They show that diversity of layers in such resource-constrained models yield better trade-offs between accuracy and latency of the model. They have shown results confirming better performance of their model then other models like MobileNet [1], SqueezeNext [16], ShuffleNet [23], MobileNetV2 [7], NASNet [30] and many other models.

Figure 1 shows the architecture of MnasNet. It uses Convolution Ops, depthwise separable convolution, mobile inverted bottleneck layers to extract features. It uses RMSProp optimizer [31], Batch Normalization [27] and Dropout regularization [32]. Table I shows the MnasNet architecture which is trained with CIFAR-10 dataset.

MnasNet Architecture					
Layers	Convolutions	t	c	n	s
$32^2 \times 3$	Conv2d 3×3	-	32	1	1
$112^2 \times 32$	SepConv 3×3	1	16	1	2
$112^2 \times 16$	MBConv3 3×3	3	24	3	2
$56^2 \times 24$	MBConv3 5×5	3	40	3	2
$28^2 \times 40$	MBConv6 5×5	6	80	3	2
$14^2 \times 80$	MBConv6 3×3	6	96	2	1
$14^2 \times 96$	MBConv6 5×5	6	192	4	1
$7^2 \times 192$	MBConv6 3×3	6	320	1	1
$7^2 \times 320$	FC,Pooling			10	

TABLE I
WHERE T: EXPANSION FACTOR, C: NUMBER OF OUTPUT CHANNELS, N: NUMBER OF BLOCKS AND S: STRIDE

III. R-MnasNet Architecture

This section will explain the proposed R-MnasNet architecture and the modifications made to the baseline architecture (MnasNet) to reduce the model size and increase the accuracy. Table II shows the R-MnasNet architecture.

A. Convolution Layers

To increase the accuracy of the model, the convolution layers must extract features in an efficient way. MnasNet uses Depthwise Separable Convolutions to extract features. Depthwise separable convolutions have two layers of convolution. First, a depthwise convolution layer which extracts the features along the spatial dimensions of the input tensor. Second, a 1x1 pointwise convolution layer which covers the depth of the input tensor. In these convolutional layer, the spatial dimensions remain constant.

In order to extract features more efficiently, new convolutional layers were added in the baseline model. These layers are known as Harmonious Bottleneck Layers [2]. Along with channel transformations, these layers take spatial dimensions into consideration as well. They change the spatial feature scale along composite layers of the network. As a result, the balance between capability of representation and computations is improved.

R-MnasNet Architecture					
Layers	Convolutions	t	c	n	s
$32^2 \times 3$	Conv2d 3×3	-	32	1	1
$112^2 \times 32$	SepConv 3×3	1	16	1	2
$112^2 \times 16$	MBConv3 3×3	3	24	3	2
$112^2 \times 4$	Harmonious Bottleneck	2	36	1	1
$56^2 \times 36$	MBConv3 5×5	3	40	3	2
$112^2 \times 40$	Harmonious Bottleneck	2	72	1	2
$28^2 \times 72$	MBConv6 5×5	6	80	3	2
$112^2 \times 80$	Harmonious Bottleneck	2	96	4	2
$14^2 \times 96$	MBConv6 3×3	6	96	2	1
$112^2 \times 80$	Harmonious Bottleneck	2	192	1	2
$112^2 \times 80$	Harmonious Bottleneck	2	96	4	2
$14^2 \times 96$	MBConv6 5×5	6	192	4	1
$112^2 \times 80$	Harmonious Bottleneck	2	288	1	1
$7^2 \times 192$	MBConv6 3×3	6	320	1	1
$7^2 \times 320$	FC,Pooling			10	

TABLE II
WHERE T: EXPANSION FACTOR, C: NUMBER OF OUTPUT CHANNELS, N: NUMBER OF BLOCKS AND S: STRIDE

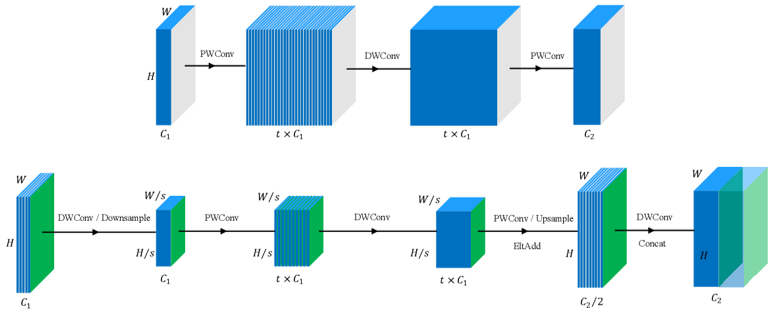


Fig. 2. Comparison of Depthwise Separable Convolution Layer and Harmonious Bottleneck Layer. [2]

As shown in Figure 2, there is contraction-expansion of spatial dimensions and expansion-contraction of channel dimensions in the harmonious bottleneck layer. First, the channel dimensions remain constant and features get extracted from the spatial dimensions. Second, the spatial dimensions are kept constant and features are extracted from channel dimensions. This makes the convolution layer more efficient and increases the accuracy of the model. It also results in decreasing the computational cost of the model. The model size decreases and makes the model lightweight.

Figure 2 shows the comparison of depthwise separable convolutional block and harmonious bottleneck layer. The spatial size of input/output feature maps is (H x W), C1/C2 are input/output feature channels, (K x K) is the kernel size and s denotes stride.
The total cost of depthwise separable convolution is

$$(H \times W \times C1 \times K \times K) + (H \times W \times C1 \times C2) \quad (1)$$

The total cost of harmonious bottleneck layer is

$$B/s^2 + (H/s \times W/s \times C1 + H \times W \times C2) \times K^2 \quad (2)$$

where, B is the computational cost of the blocks inserted between the spatial contraction and expansion operations. It is evident that by squeezing the channel expansion-contraction component and using a pair of spatial transformations yields a slimmed spatial size of wide feature maps in each stage which reduces the computational cost.

In this paper, six harmonious bottleneck layers were added to the MnasNet architecture. The model is described in table II. After adding these layers, the accuracy of the model increased by 9.34% whereas the model size decreased from 12.7 MB to 3 MB. The comparison is shown in Table III.

B. Learning Rate Annealing or Scheduling

Different learning rates are used while training the network. There are various methods to choose the learning rate for particular epochs during training. Time-based decay, step decay and exponential decay are some of the methods which are used to determine learning rates during training. Figure 3 illustrates that step decay based learning rate performs better than other learning rate schedule methods. Adaptive learning rate methods are also used instead of manually scheduling the learning rates. In this paper, the step decay method proved to be more efficient. Therefore, this method is used for training R-MnasNet.

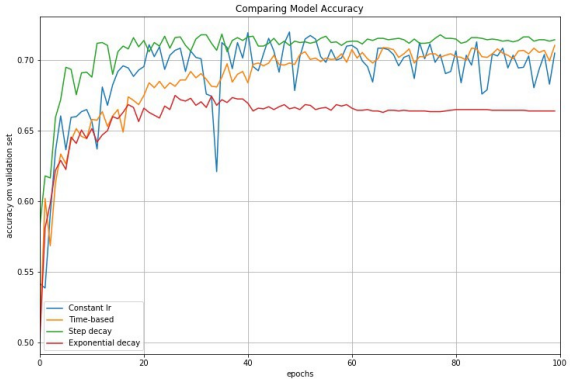


Fig. 3. Comparision of different LR scheduling methods.

C. Optimization

There are various optimization techniques which are used to optimize CNNs. Some commonly used optimizers are Adaptive Learning Rate Method (Adadelta) [31], Adaptive Gradient Learning Algorithm (Adagrad) [31], Adaptive Moment Estimation (Adam) [31], Stochastic Gradient Descent (SGD) [31], Root Mean Square Propagation (RMSprop). RMSprop was used to optimize the MnasNet architecture. In this paper, we have used SGD with varying values of learning rates and momentum=0.9 for better optimization.

D. Data Augmentation

AutoAugment [3] was used for data augmentation. AutoAugment learns the best augmentation policies for a given dataset with the help of Reinforcement Learning (RL). A policy consists of 5 sub-policies and each sub-policy applies 2 image operations in sequence. Each of those image operations has two parameters: The probability of applying it and the magnitude of the operation (e.g. rotate 20 degrees in 65% of cases). There is a controller that decides the best data augmentation policy at that instant and tests the generalization ability of that policy by running a child model experiment on a small subset of a particular dataset. After the child experiment is finished the controller is updated with the validation accuracy as the reward signal, using a policy gradient method called Proximal Policy Optimization algorithm (PPO). In this paper, AutoAugment is used on CIFAR-10 dataset. The accuracy of R-MnasNet was 88.54% but after using AutoAugment the accuracy increased to 90.14%.

E. Mish Activation Function

Activation functions are used to introduce non-linearity in neural networks. They determine the correct non-linear relation between the input and output signals. In 2019, Mish [35] was introduced and it outperformed all other activation functions. It is a new type of gated softplus function. The softplus activation function can be represented as:

$$\varsigma(x) = \ln(1 + x) \quad (3)$$

Mathematically, Mish can be written as,

$$f(x) = x \cdot \tanh(\varsigma(x)) \quad (4)$$

Figure 6 shows the graphical representation of Mish. For comparison, Figure 7 shows commonly used activation functions along with the graph of Mish activation.

Mish [35] avoids saturation due to near zero gradients, strong regularization effects, preserves small negative gradients and has effective optimization and generalization. After implementing it in R-MnasNet, the acurracy of the model increased from 90.14% to 91.13%.

IV. Hardware and Software used

- Intel i9 9th generation processor with 32 GB RAM
- Aorus Geforce RTX 2080Ti GPU
- Python version 3.6.7.
- Pytorch version 1.0.
- Spyder version 3.6.
- Livelossplot

V. Results

The accuracy of baseline architecture is 80.8% when trained with CIFAR-10 [4] dataset with a model size of 12.7 MB. After introducing new layers, the accuracy became 88.84% and the model size reduced to 3 MB. AutoAugment was implemented to further incerase the accuracy. It was evident that the accuracy increased from 88.54% to 90.14%. Furthermore,

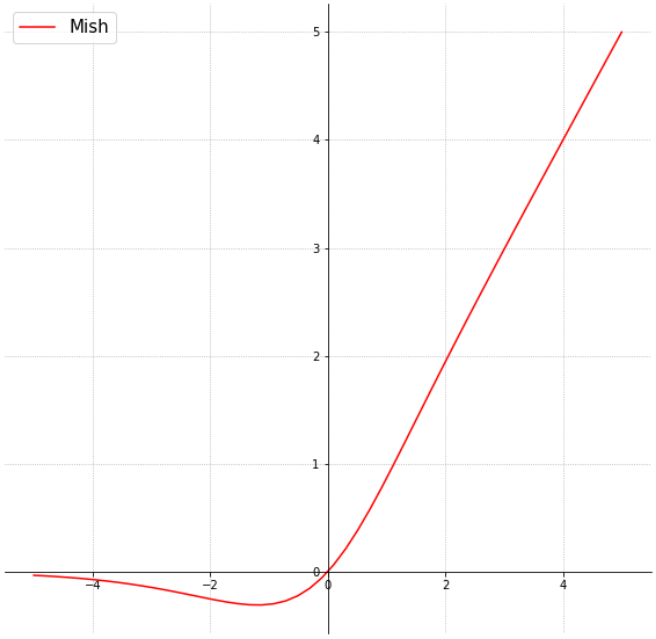


Fig. 4. Mish Activation Function

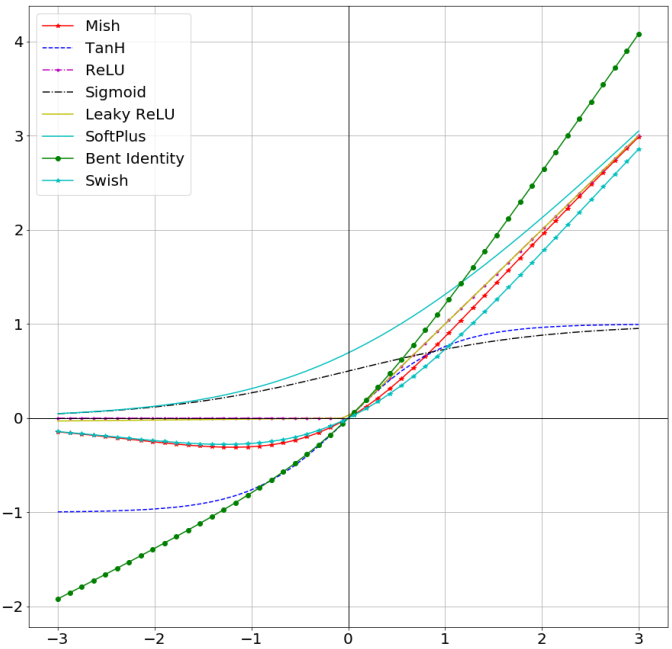


Fig. 5. Common Activation Functions

ReLU6 activation function was replaced with Mish activation function. This resulted with the model accuracy of 91.13%.

The aforementioned modifications were made to the baseline architecture. The model was trained with CIFAR-10 [4] dataset on Aorus Geforce RTX 2080Ti GPU using PyTorch framework for 200 epochs. The data was divided into batch size of 128 for training set and batch size of 64 for validation set.

Figure 6 and figure 7 shows the plots of log loss and accuracy of their respective models. The graphs were plotted using livelossplot visualization library.

The trade-off between model size and accuracy is shown in table III

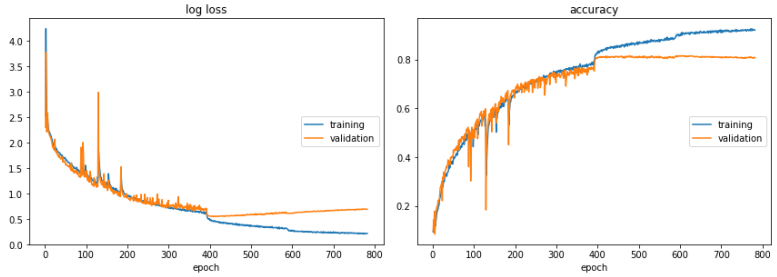


Fig. 6. Baseline

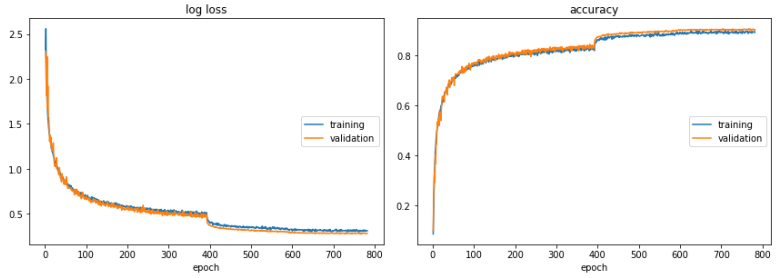


Fig. 7. R-MnasNet

Comparison of models		
Architecture	Model Accuracy	Model size (in MB)
MnasNet	80.8%	12.7
R-MnasNet	91.13%	3

TABLE III

Table IV shows the results obtained by scaling the model with different values of width multiplier.

Scaling R-MnasNet with width multiplier		
Width Multiplier	Model Accuracy	Model size
1.4	92.49%	5.6 MB
1.0	91.13%	3 MB
0.75	90.03%	2 MB
0.5	87.5%	1.3 MB
0.35	84.9%	837.6 KB

TABLE IV

VI. Conclusion

In this paper, a new architecture, R-MnasNet, is introduced by modifying its baseline architecture. The goal was to make the model more compact and having a fair trade-off between model size and accuracy. The accuracy of R-MnasNet is 91.14% with a size of 3 MB. It outperforms its baseline architecture MnasNet which has an accuracy of 80.8% and model

size of 12.6 MB. New Harmonious Bottleneck layers were added to the baseline architecture. Mish activation was used to improve the optimization of the network. AutoAugment was used to further increase the accuracy of the model. This model can be deployed on mobile devices and various embedded systems. It can be used for embedded vision applications. Depending on the application and hardware, a particular variant of R-MnasNet can be deployed by scaling it with width multiplier to achieve specific model size and accuracy.

REFERENCES

[1] Mingxing Tan, Bo Chen, et al. "MnasNet: Platform-Aware Neural Architecture Search for Mobile" arXiv:1807.11626v3 [cs.CV] 29 May 2019

[2] Duo Li, Aojun Zhou, Anbang Yao. "HBONet: Harmonious Bottleneck on Two Orthogonal Dimensions" arXiv:1908.03888v1 [cs.CV] 11 August 2019

[3] Ekin D. Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, QuocV. Le Google Brain, "AutoAugment: Learning Augmentation Strategies from Data" arXiv:1805.09501v3 (2019)

[4] https://www.cs.toronto.edu/ kriz/cifar.html

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, " Imagenet classification with deep convolutional neural networks" In Advances in neural information processing systems, pages 10971105, 2012.

[6] Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." arXiv:1704.04861 (2017).

[7] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, Liang-Chieh Chen. "MobileNetV2: Inverted Residuals and Linear Bottlenecks." arXiv:1801.04381v4 (2019)

[8] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z.Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. "Imagenet largescale visual recognition challenge", International Journal of Computer Vision, 2015.

[9] J. Wu, C. Leng, Y. Wang, Q. Hu, and J. Cheng. "Quantized convolutionalneural networks for mobile devices". arXiv preprint arXiv:1512.06473,2015.

[10] W. Chen, J. T. Wilson, S. Tyree, K. Q. Weinberger, and Y.Chen. "Compressing neural networks with the hashing trick". CoRR,abs/1504.04788, 2015

[11] S. Han, H. Mao, and W. J. Dally."Deep compression: Compressing deep neural network with pruning, trained quantization and huffman coding".CoRR, abs/1510.00149, 2, 2015.

[12] Karen Simonyan, Andrew Zisserman."VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION" arXiv preprint arXiv:1409.1556v6 (2015)

[13] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. "Rethinking the inception architecture for computer vision." arXiv preprint arXiv:1512.00567, 2015.

[14] C. Szegedy, S. Ioffe, and V. Vanhoucke. Inception-v4, inception-resnet and the impact of residual connections on learning. arXiv preprint-arXiv:1602.07261, 2016

[15] Sebastian Ruder. "An overview of gradient descent optimization algorithms." arXiv preprint arXiv:1609.04747, 2017

[16] F. N. Iandola, M. W. Moskewicz, K. Ashraf, S. Han, W. J. Dally, and K.Keutzer. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 1MB model size. arXiv preprint arXiv:1602.07360, 2016.

[17] Amir Gholami, Kiseok Kwon, Bichen Wu, Zizheng Tai, Xiangyu Yue, Peter Jin, Sicheng Zhao, Kurt Keutzer "SqueezeNext: Hardware-Aware Neural Network Design." arXiv preprint arXiv:1803.10615v2

[18] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun "Deep Residual Learning for Image Recognition." arXiv preprint arXiv: 1512.03385 (2015).

[19] Chen Wang, Yang Xi. Convolutional Neural Network for Image Classification.

[20] MD. ZAKIR HOSSAIN, FERDOUS SOHEL, MOHD FAIRUZ SHIRATUDDIN, HAMID LAGA (2018). "A Comprehensive Survey of Deep Learning for Image Captioning." arXiv preprint arXiv: 1810.04020.

[21] Zhong-Qiu Zhao, Shou-tao Xu, and Xindong Wu. "Object Detection with Deep Learning: A Review." arXiv preprint arXiv:1807.05511 (2019).

[22] Abien Fred M. Agarap. "Deep Learning using Rectified Linear Units (ReLU)." arXiv preprint arXiv:1803.08375v2[cs.NE], 2019

[23] Xiangyu Zhang, Xinyu Zhou, Mengxiao LinJian, Sun. "ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile." arXiv preprint arXiv:1707.01083v2, 2017

[24] Francois Chollet (2017). "Xception: Deep Learning with Depthwise Separable Convolutions. arXiv preprint arXiv:1610.02357

[25] Zhun Zhong, Liang Zheng, Guoliang Kang, Shaozi Li, Yi Yang (2017). "Random Erasing Data Augmentation." arXiv preprint arXiv:1708.04896

[26] Barret Zoph, Ekin D. Cubuk, Golnaz Ghiasi, Tsung-Yi Lin, Jonathon Shlens, Quoc V (2019). Le. "Learning Data Augmentation" arXiv preprint arXiv: 1906.11172

[27] Sergey Ioffe and Christian Szegedy. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift." arXiv preprint arXiv: 1502.03167, 2015

[28] Timothy Dozat (2016). "INCORPORATING NESTEROV MOMENTUMINTO ADAM." Workshop track-ICLR 2016.

[29] Diederik Kingma and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980, 2014

[30] Barret Zoph, Vijay Vasudevan, Jonathon Shlens and Quoc V. Le, "Learning Transferable Architectures for Scalable Image Recognition", arXiv:1707.07012v4 [cs.CV] 11 Apr 2018.

[31] Sebastian Ruder, "An overview of gradient descent optimization algorithms*",arXiv:1609.04747v2 [cs.LG] 15 Jun 2017

[32] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting", Journal of Machine Learning Research 15 (2014) 1929-1958, Submitted 11/13; Published 6/14.

[33] https://towardsdatascience.com/learning-rate-schedules-and-adaptive-learning-rate-methods-for-deep-learning-2c8f433990d1

[34] http://www.image-net.org/

[35] Diganta Misra, "Mish: A Self Regularized Non-Monotonic Neural Activation Function", arXiv preprint arxiv:1908.08681 (2019)

IoT Antennas for Industry 4.0 – Design and Manufacturing with an Example

Prabakar Parthiban
Department of Engineering
Auckland University of Technology
New Zealand
prabakarparthiban@gmail.com

***Abstract*— Automation and Internet-of-Things (IoT) open doors to achieve increased productivity and efficiency. Ultra-High Frequency (UHF) Radio-Frequency Identification (RFID) is a type of Auto-identification technique that identifies, locates, authenticates, and engages with other IoT devices to efficiently perform one or more tasks. The concept of the digital twin is enhanced by UHF RFID for the fourth-generation industrial revolution, Industry 4.0. UHF RFID antennas are one of the key endpoint components that are deployed throughout factories, facilitating automation. Antennas predominantly used in industrial environment requires enormous durability to handle vibration, shock, impact, and thermal shocks without trading-off the radiofrequency (RF) performance. This article proposes a novel durable dual-virtual patch antenna design and manufacturing technique for industrial UHF RFID application. The durability of the antenna is offered by the non-conventional RF substrate, galvanized steel backplate and the polyurethane resin-filled plastic radome. The antenna is designed to be assembled by full and semi-robotic process automation.**

Keywords*— *IoT antenna, RFID antenna, industry 4.0, rugged antenna, ETSI.

I. INTRODUCTION

The fourth-generation industrial revolution (industry 4.0 or I4.0) demands digitization as a core component in the manufacturing sector. Advancements in sensing and communication technologies has leveraged high visibility in the manufacturing processes. Internet of things (IoT) enabled automation helps to gather information of products, processes, and stages through sensors, real-time. This information is known as the digital twin – with which manufacturing processes can be managed efficiently with high transparency, efficiency, accuracy, and reduced risks. IoT is an umbrella of various embedded computer systems, sensors such as the wireless sensor networks and control systems and automation where automatic identification and data-capture (AIDC) is an integral part of it. Some common types of AIDC sensors are, Optical sensors such as barcodes, digital sensors such as cameras and wireless sensors such as radio-frequency identification (RFID). Passive RFID technology such as Ultra-high frequency (UHF) makes the I4.0's dream viable as it can identify, locate, authenticate and engages an everyday item with other IoT devices. An industry alliance called the RAIN RFID (Radio frequency IdentificatioN) is promoting world-wide adoption of UHF RFID in a similar way to other alliances such as the WiFi alliance, Bluetooth SIG, etc.

A typical passive UHF RFID system consists of a transceiver commonly known as a reader, reader antenna which transmits and receives radio-frequency signals from a passive RFID tag (refer Figure 1). The passive RFID tag consists of a tiny chip attached to a printed antenna which receives the energy and backscatters to the reader antenna. This tag can be attached to almost all the everyday items around us. Reader antennas are distributed across hallways, entrance-ways or warehouses to track the assets' movements. Reader antennas are energized by one or more readers who communicate the received tag data over Wi-Fi or Bluetooth to a server. As these reader antennas are prone to knocks and whacks (due to forklift movements, etc.), and it is necessary to have antennas that are rugged and durable. Durable antennas can withstand vibrations, mechanical shocks, impacts and thermal shocks. In general, RFID tags are made durable to last long throughout the supply-chain process, but reader antennas require extra protection such as an enclosure to handle knocks and whacks within the factory. This article proposes a durable antenna design and manufacturing technique that does not require any enclosure during installation.

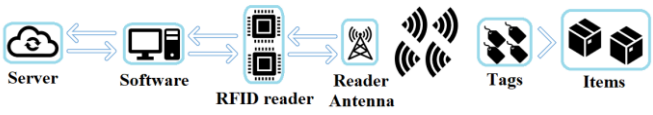


Fig. 1. A typical UHF/RAIN RFID system

II. Literature Review

Literature in RFID antenna design reports various durable tag antenna designs such as [1-4], but no RFID reader antennas are evident. Thus, durable and rugged antenna designs used for other applications are used for comparison and review. Traditionally, expensive, rugged substrates such as Arlon AD255A [5] are used for durable antenna designs. The proposed substrate is low-cost and widely available with low dielectric constant (ε_r), 1.5 unlike Arlon's 2.55. The proposed antenna's durability is also enhanced by its rugged housing construction which is lacking in [5]. The proposed antenna design has a directional radiation pattern to focus the antenna beam in the desired direction as opposed to the monopole antenna design in [5], whose radiation is not focused but omnidirectional. Reader antennas with

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

unfocussed radiation pattern are not suitable for UHF RFID applications as they would detect unwanted items (often referred to as stray reads) [6]. The mechanically robust antenna design shown in [7] is based on a cross-slot antenna array on a rugged metal cylinder without a protective radome. This antenna is designed for higher frequencies such as 2.5 GHz and not suitable for UHF RFID applications. Besides, its omnidirectional radiation creates more stray reads like [5].

The ultra-rugged UHF RFID monopole antenna affixed on a concrete substrate with a ground plane is proposed in [8]. The antenna's radiation is focused in one direction using the reflective ground plane, unlike the monopole in [5]. The antenna does not have a rugged protective radome but relies on the substrate's rigidity. The radome, substrate and the stainless-steel ground plane collectively offers the durability in the proposed antenna. For instance, when the antenna in [8] is subject to a mechanical impact, the monopole antenna strip and the cable solder joint can be badly damaged, leaving the antenna impaired. The radome in the proposed antenna can withstand impacts and leave the antenna unimpaired. Furthermore, the cable is protected integrally within the antenna, and thus the solder joint cannot be fractured in contrast to [8]. Other embeddable patch antenna designs shown in [9-11] relies on the concrete cover-like structure for protection as the design lacks a durable radome. In other words, these antennas are not durable when they are not embedded in concrete. The proposed antenna is durable on its own with its ruggedized protective radome. The antenna, including the radome design and manufacturing process, is discussed in this paper, whereas [9-11] did not explain in detail about the manufacturing process but focused only on the design.

III. Antenna Design

A durable dual-virtual patch antenna array on a foamed PVC substrate and galvanized steel ground plane is designed for industrial applications. The patch antenna array made from tinned steel. The foamed PVC substrate is a non-conventional substrate that is not used for RF applications. PVC substrate is previously used in [12 and 13] due to its inherited properties such as low moisture absorption, high chemical resistance, self-extinguishing flammability rating, easy to fabricate and light-weight. Foamed PVC substrate is chosen for the same purposes. The substrate's thickness is 6 mm, and it is 220 mm wide and long. The thickness of the galvanized steel ground plane is 2 mm. The ground plane is square in shape and spans 220 x 220 mm. Steel is highly strong, and galvanization helps to avoid rusting. A 6 mm 'L' shaped groove shown in Fig. 1 is implemented to recess the antenna cable flush on the ground plane. The length and width of each of the patch in the array is 124 mm and 88 mm, respectively for operation in 864-868 MHz – European RFID frequencies. Patches are not spaced for asymmetric beamforming but are connected together to obtain a symmetric radiation pattern. Patches are inset fed using 100 Ω microstrip lines at the 100 Ω impedance points. The 50 Ω input impedance is divided into two 100 Ω microstrip lines at the 'T' junction. A 50 Ω RG-58 coaxial cable with N-type connector is soldered at the 'T' junction. A 30 x 20 mm ground patch is located right next to the cable groove to solder the coaxial cable's ground. N-type connectors rugged compared to TNC or SMA connectors. Tin plated steel is rust-proof and enables ease of soldering. The thickness of the tinned steel patch antenna array element is 0.2 mm. The ground patch is referenced to the antenna's ground plane using solid vias. Aluminum rivets (with 6 mm diameter) are used as vias that are attached between the ground patch and the ground plane. The antenna would include a protective radome. The radome comprises of a formed plastic tray made from Kydex-100 thermoplastic that is filled with polyurethane resin. The tray is 241 mm long, 241 mm wide and 15 mm deep and the resin is filled until 6 mm deep. The radome is made only slightly larger so that the antenna fits snug. The following section explains the manufacturing process and fabrication.

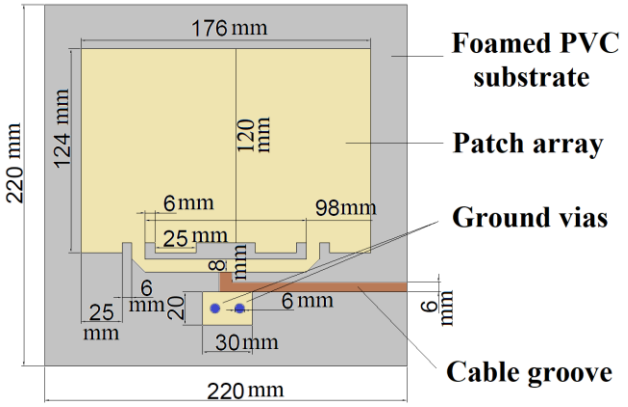


Fig. 1. Durable antenna design

IV. Manufacturing Process and Fabrication

Different manufacturing techniques are adopted to manufacture the antenna parts and fabricate the antenna. The radiating element (patch antenna array and the microstrip line), the ground plane with mounting holes, ground patch with via holes and the grooved substrate are laser cut. The ABS plastic tray radome is vacuum- formed. As these techniques are non-exotic, antenna parts can be sourced from machine shops. The manufacturing process outlined in Fig. 2 begins with sourcing all the parts viz., ground plane, patch antenna array, ground patch, grooved substrate and vacuum-formed plastic. The next step in the process is to source the consumables such as rivets, solder, adhesive and polyurethane resin. Sourced parts undergo a quality check process to ensure that the parts meet specifications. Only flawless parts are put together to fabricate the antenna. Damaged parts or parts that do not meet the specifications are returned and resourced.

The substrate is attached to the ground plane using the acrylic adhesive. The patch antenna and the ground patch are attached to the substrate, just above and below the groove,

respectively, using the acrylic adhesive. These attachment process can be robotized or labored using placement jigs to maintain reliability and repeatability. The coaxial cable is slid inside the groove to solder its center conductor with the microstrip 'T' junction and the ground braid with the ground patch (See Fig. 3(a)). Soldering process can be robotized to obtain repeatable solder joints as it is critical for proper impedance matching. Rivets are popped using pneumatic or electric rivet guns to make reliable connections between the ground patch and the galvanized steel ground plane (See Fig. 3(c)). Gaps in the cable groove are sealed with a two-part epoxy resin that is chemical and heat resistant. Cable strain relief is provided by this epoxy seal. Sealing process can also be automated using robots or labored using jigs.

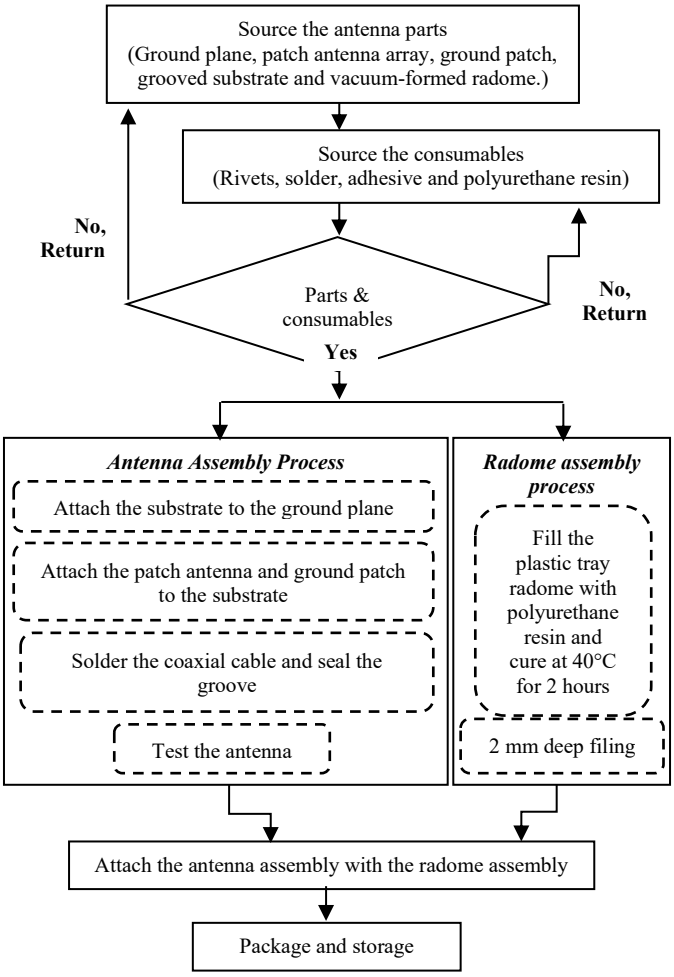


Fig. 2. Manufacturing flow diagram

The plastic tray is filled with polyurethane resin using a fixed nozzle dripper that has a constant rate of flow and a time-controlled stopper. The resin consists of two parts – resin and hardener. The resin cures in 2 hours at 40 °C to form a 6 mm thick layer. The flow rate, nozzle diameter, and stopping time is critical to achieving the 6 mm layer thickness. The antenna is tuned to this layer thickness, and thus any changes to the thickness lead to antenna detuning. The flow rate can be controlled by a pneumatic or electrical syringe, and the stopper can be clocked using a micro-controller. The cured radome is filed using a linisher to remove 2 mm polyurethane material. Filing is required to avoid the stress caused in the solder joint when adhering the radome (see Fig. 3(b)). The antenna is tested for its standing wave ratio (SWR) at 866.5 MHz (center of the operating band) using an SWR meter to ensure proper solder connections and impedance matching.

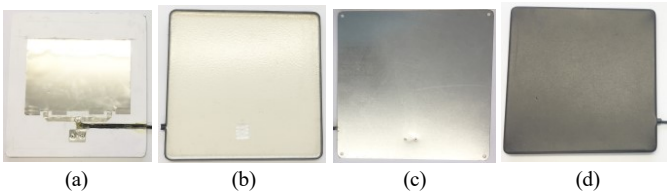


Fig. 3. (a) Fabricated antenna; (b) Radome; (c) Ground plane; and (d) Fully assembled durable antenna

The antenna assembly and radome assembly can happen in parallel, as there are no dependencies within these processes. Fully assembled antenna and radome are finally attached together using an acrylic adhesive. The fully assembled durable antenna is shown in Fig. 3(d). Table. I list the antenna's Bill of Materials (BoM). The total parts' cost is no more than $14.63 (USD), excluding labor. It is assumed that parts such as the substrate, ground plane, radiating array and the vacuum-formed radome are obtained from suppliers and thus, there are no capital costs to be reported. Labor is estimated at $2.66 based on the federal minimum wage, $7.25 (USD) per hour. The total cost for manufacturing per antenna is $17.30 (USD). The durable antenna acquires the inherited physical properties of the ABS cover, polyurethane resin, foamed PVS substrate and the steel ground plane. Table II lists the antenna's physical properties.

TABLE I
DURABLE ANTENNA'S BILL OF MATERIALS

Parts	Cost (USD)	Quantity used	Sub-total (USD)
ABS sheet	$45 for a sheet (1290 x 2540 mm)	300 x 300 mm	$1.236
Polyurethane resin	$17.3 per 250g	100g	$6.92
Tin sheet radiator	$35 for a sheet (1200 x 2400 mm)	175 x 130 mm	$0.27
RG-58 Cable	$3 for cable assembly	1.0	$3.00
Foamed PVC	$50 for a sheet (1200 x 2400 mm)	220 x 220 mm	$0.84
Galvanized steel	$1.80 for 220 x 200 mm sheet	1.0	$1.80
Rivets	$0.05	2.0	$0.10
Acrylic adhesive	$23 for 36 mm x 50000 mm reel	36 mm x 1000 mm	$0.46
		Total Bill of Materials (BoM):	**$14.63**

TABLE II
DURABLE ANTENNA'S PHYSICAL PROPERTIES

Parameter	Durable antenna
Tensile strength and modulus	51 MPa and 2320 MPa
Compression modulus and strength	1910 MPa and 64 MPa
Flexural strength	79 MPa

V. Measurements and Testing

An anechoic chamber is used to measure high-frequency radiating devices such as antennas [14]. Return loss is the measures using TR-1300/1vector network analyzer (VNA). The return loss of less than -10 dB is required for an efficient RFID performance. The antenna's return loss is around -12 dB (better than the required threshold) in 864 to 868 MHz (see Fig. 4). The antenna's -10 dB return loss bandwidth is 22 MHz. The antenna's gain is measured by the gain-comparison method [15]. The gain is flat across the frequencies of interest (865-868 MHz) and measures 9.9 dBi. The antenna's radiation characteristics are also measured in azimuth and elevation cut planes. The antenna's half-power beam-width (HPBW) is 60° in both azimuth and elevation planes. The front-to-back ratio (FBR) is measured to be -12 dB. There are no side lobes in this array antenna because the patches are not separate but virtually unconnected. The radiation pattern of the antenna is shown in Fig. 5.

Temperature and relative humidity (RH) inside a factory or an industrial plant can range somewhere between 0° and 50° & 45% and 85%, respectively, depending on their geographical location. A range beyond this may be unsafe for people, and thus this assumed optimal range is used for test purposes. It was noted that varying temperature at 85% RH yielded noticeable changes in the antenna's performance compared to lower RH settings. The antenna's return loss for 0°C to 50°C temperature range at 85% RH is plotted in Fig. 6. As return loss and gain are directly proportional (based on Fig. 4), only return loss values are measured to confirm the antenna's operational efficiency.

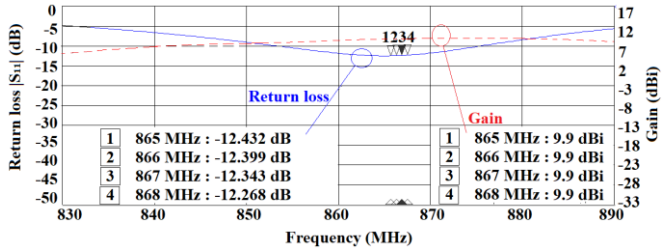


Fig. 4. Return loss and Gain measurements

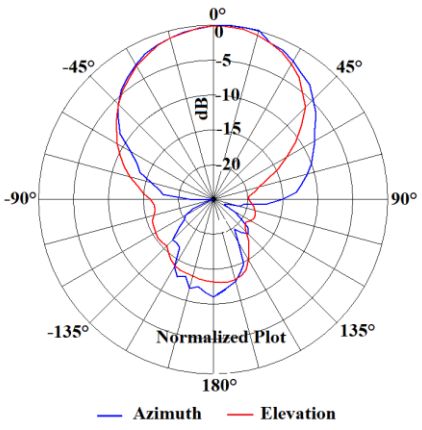


Fig. 5. Radiation Pattern

When the temperature is higher, the antenna tends to resonate at a slightly lower frequency and vice versa for lower temperature ranges such as 10°C and 0°C. Nevertheless, return loss values at the frequencies of interest (865-868 MHz) are less than -10 dB. The antenna's performance is not degraded with the change in temperature and humidity. The antenna is also tested with large metal backings to see its insusceptibility as these antennas are usually mounted on industrial equipment and machinery. The antenna's resonance is almost unaltered with both 300 x 600 mm and 600 x 600 mm sized metal reflector backings.

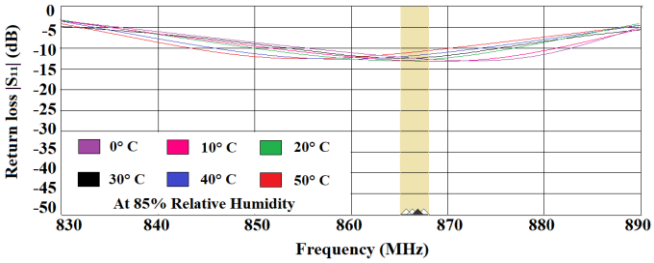


Fig. 6. Effect of temperature and humidity

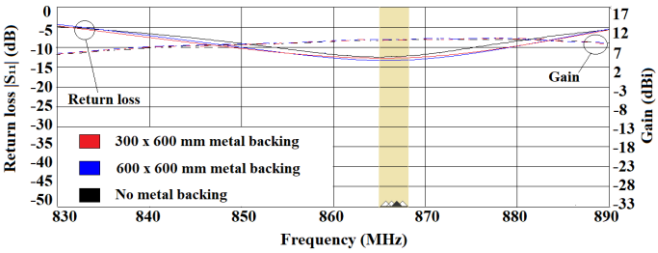


Fig. 7. Effect of metal backing

VI. Conclusion and Future Work

A durable UHF RFID reader antenna design and manufacturing technique for industry 4.0 are proposed. The durability of the antenna is offered by the non-conventional RF substrate, galvanized steel back plate and the polyurethane resin-filled plastic radome. The antenna is linearly polarized and has 20 MHz operating bandwidth with 9.9 dBi gain. The antenna's radiation pattern is symmetrical in both planes, thus increasing the tag read efficiency. The antenna is rugged both physically and electrically to environmental impacts such as temperature and humidity shocks, physical impacts, vibrations, and effect of different metal back plates. As part of the future work, the antenna will be tested for its real-life RFID performance in exemplar scenarios.

Acknowledgment

The author would like to thank Auckland University of Technology, New Zealand for the help with journal database and library access. I want to extend my thanks to Times-7 Research Ltd for the motivation and support in my research interests and fabricating the prototype.

VII. REFERENCES

[1] M. Toivonen, T. Björninen, L. Sydänheimo, L. Ukkonen and Y. Rahmat-Samii, "Impact of Moisture and Washing on the Performance of Embroidered UHF RFID Tags," *IEEE Antennas and Wireless Propagation Letters*, vol. 12, pp. 1590-1593, 2013.

[2] K. Koski *et al*., "Durability of embroidered antennas in wireless body-centric healthcare applications," *2013 7th European Conference on Antennas and Propagation (EuCAP)*, Gothenburg, 2013, pp. 565-569.

[3] Y. Y. Fu *et al*., "Experimental Study on the Washing Durability of Electro-Textile UHF RFID Tags," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 466-469, 2015.

[4] T. Phatarachaisakul, T. Pumpoung, C. Phongcharoenpanich and S. Kosulvit, "Tag antenna using printed dipole with H-slot for UHF RFID applications," *2014 International Electrical Engineering Congress (iEECON)*, Chonburi, 2014, pp. 1-4.

[5] Ting Wu, Xiao-Wei Shi, Ping Li and Hao Bai, "Tri-band microstrip-fed monopole antenna with dual-polarisation characteristics for WLAN and WiMAX applications," *Electronics Letters*, vol. 49, no. 25, pp. 1597 -1598, 5 December 2013.

[6] P. Parthiban, "Fixed UHF RFID Reader Antenna Design for Practical Applications: A Guide for Antenna Engineers With Examples," IEEE Journal of Radio Frequency Identification, vol. 3, no. 3, pp. 191-204, Sept. 2019

[7] M. S. Sadiq and C. Ruan, "A Compact, Mechanically rugged, DC Grounded 45° Slant Polarized Low Gain Ripple Omnidirectional Antenna," 2019 IEEE MTT-S International Wireless Symposium (IWS), Guangzhou, China, 2019, pp. 1-3.

[8] P. Parthiban, B. Seet and X. J. Li, "An Ultra-Rugged UHF RFID Reader Antenna: A New Design for Convention Center Loading Docks [Wireless Corner]," IEEE Antennas and Propagation Magazine, vol. 62, no. 1, pp. 84-95, Feb. 2020.

[9] R. Salama and S. Kharkovsky, "An embeddable microwave patch antenna module for civil engineering applications," 2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Minneapolis, MN, 2013, pp. 27-30.

[10] S. Jeong and H. Son, "UHF RFID Tag Antenna for Embedded Use in a Concrete Floor," IEEE Antennas and Wireless Propagation Letters, vol. 10, pp. 1158-1161, 2011.

[11] K. M. Z. Shams and M. Ali, "Wireless Power Transmission to a Buried Sensor in Concrete," IEEE Sensors Journal, vol. 7, no. 12, pp. 1573-1577, Dec. 2007.

[12] P. Parthiban, B. Seet and Xue Jun Li, "Low-cost scalable UHF RFID reader antenna with no surface dead zones," *2016 IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE)*, Langkawi, 2016, pp. 24-29.

[13] P. Parthiban, B. Seet and X. J. Li, "Display Frame UHF RFID Antenna for Museums," *2018 IEEE Asia-Pacific Conference on Antennas and Propagation (APCAP)*, Auckland, 2018, pp. 532-533.

[14] I. Munteanu and R. Kakerow, "Simulation Methodology for the Assessment of Field Uniformity in a Large Anechoic Chamber," IEEE Transactions on Magnetics, vol. 50, no. 2, pp. 213-216, Feb. 2014.

[15] IEEE Standard Test Procedures for Antennas," in ANSI/IEEE Std 149-1979 , vol., no., pp.0_1-, 1979

OPTIMAL MANAGEMENT OF WATER RESOURCES IN LARGE MAIN CANALS WITH CASCADE PUMPING STATIONS

Anvar Kabulov
Department of "Applied mathematics and intellectual technologies
National University of Uzbekistan named after Mirzo Ulugbek
Tashkent, Uzbekistan
anvarkabulov@mail.ru

Ibrokhimali Normatov
Department of "Applied mathematics and intellectual technologies
National University *of Uzbekistan named after Mirzo Ulugbek*
Tashkent, Uzbekistan
ibragim_normatov@mail.ru

Aybek Seytov
Department of "Mathematics"
Chirchik State Pedagogical Institute
Chirchik, Uzbekistan
seytov.aybek@mail.ru

Adilbay Kudaybergenov
def. of Applied mathematics
Karakalpak State University named after Berdakh
Nukus, Uzbekistan
q_adilbay@karsu.uz

***Abstract*—A mathematical model of the unsteady flow of water on sections of main canals by limited hydraulic structures has been developed. And also the statement of problems of optimal control of water distribution in the main canals is developed. A fundamental solution is obtained for a one-dimensional differential equation describing the unsteady movement of water in sections of main canals.**

Keywords— mathematical model, unsteady flow of water, main canals, optimal control problems, fundamental solution, differential equation, hydraulic structures

I. INTRODUCTION

The development of computerization around the world, and in recent years in the Central Asian states, makes it possible to use it for modeling and managing water resources in all water-management systems and facilities, which include river sections, reservoirs and canals (free-flow and with water-lifting engine systems). Using a computer, the behavior of individual waterworks facilities, their interaction can be simulated by taking into account all the influencing factors in conditions close to real, in the process of their functioning.

Water-management systems and facilities have a large spatial extent, a multitude of technological and technical parameters, only by methods of mathematical modeling can quantitative and qualitative changes in their characteristics be obtained.

Currently, special attention is paid to saving water resources to provide consumers with them, achieving this by managing modern information systems of their operating modes on the basis of specified criteria for implementing the optimal management task.

II. METHODS AND RESULTS

We consider a model statement of the problems of water distribution in this section of the canal. As a mathematical model of unsteady water flow in the canal sections, a one-dimensional differential equation is considered in the form [1,2,3]

$$\frac{\partial \upsilon(x,t)}{\partial t}+\upsilon\frac{\partial Q(x,t)}{\partial x}=q(x,t),, \tag{1}$$

where, $Q(x,y)$ – the change in water flow in the canal section, υ – the flow rate.

Initial condition

$$Q(x,0)=Q_0(x), \tag{2}$$

where, $Q_0(x)$ – the initial distribution of water flow in the canal section.

Boundary condition

$$Q(0,t)=Q_1(t), \tag{3}$$

where, $Q_1(x)$ – the change in water flow at the beginning of the canal section.

Variable Definition Area

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

$$x \geq 0, \quad t \geq 0, \quad \upsilon > 0. \tag{4}$$

Water flows at the points of water intake of the canal section $q(x,y)$ the uniform distribution of water has the form [4,5]

$$q(x,t) = -\sum_{i=1}^{n} q_i \delta(x - a_i) 1(t - T). \tag{5}$$

Equation (5) takes into account the main properties of the irrigation canal, such as the delay in water flow along the length of the canal. The change in water flow in the initial transit of the channel section leads to its change in other transits of the channel section after a certain time, this is the delay. The farther the transit is considered from the initial transit, the greater the delay in water flow.

The fundamental solution of equation (1) - the Green function has the form [4-10]

$$G(x,\xi,t) = l(x-\xi)\delta\left(\upsilon t - (x-\xi)\right), \tag{6}$$

where , $\delta(x)$ – the Dirac delta function [6].

The analytical solution of equation (1) in the presence of its fundamental solution is determined as follows

$$Q(x,t) = \int_{t_0}^{t}\int_{0}^{l} G(x,\xi,t-\tau)\omega(\xi,t)d\xi d\tau, \tag{7}$$

where, $\omega(x,y)$ – the standardizing function for the boundary task (1) - (4), which has the form

$$\omega(x,t) = q(x,t) + Q_0(x,t)\delta(x,t)\delta(t) - \upsilon\delta(x)Q_1(t) \tag{8}$$

We consider the properties of the solution of equation (1). The change in water flow at the beginning of the canal section under zero initial conditions propagates with a speed v along the canal length. The wave front in this case does not change in length, in figure 1. it is shown the propagation of a wave of water flow along the length and time in isometry. In figure a, b in isometry the propagation of complex waves along the length of the channel is also shown.

Equation (1) takes into account the main properties of the irrigation canal, such as the delay and transformation of water flow along the length of the canal. In this case, the water flow changed in the initial transit of the canal section leads to a change in water flow at other transits of the canal section after a certain time, this is called a delay, which gradually changes in time.

In this case, the main task of water distribution is the timely interception of water flow by taps as the water flows along the length of the canal [7,8].

We consider the statement of problem of optimal control of water distribution in the canals of irrigation systems, in an example of the mathematical models developed above.

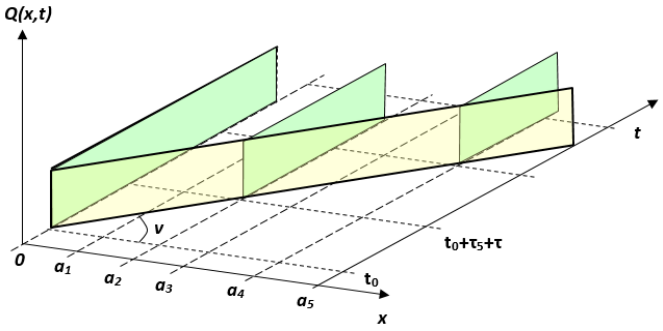
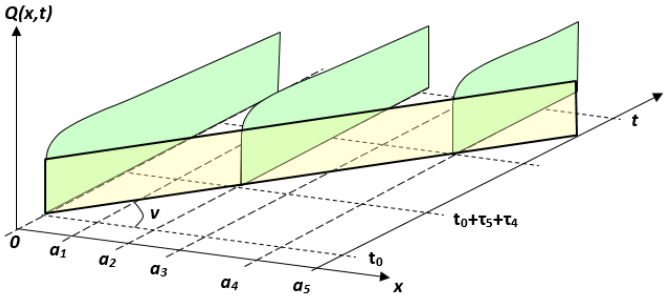
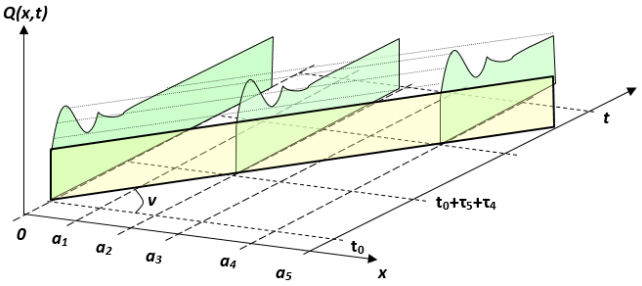


Fig. 1. The propagation of a square wave along the length of the canal



a)



b)

Fig. 2. The propagation of complex waves along the length of the canal

The main objective of the task is to minimize fluctuations in the water flow in the canal section when supplying water to the side water intakes by controlling the water flow at the beginning of the canal section, [7]

$$I = \int_{0}^{T}\int_{0}^{l}\left[Q(x,t) - Q^*\right]^2 dxdt \to min, \qquad \Omega = \left\{R^N \mid Q_{min} \leq Q(x,t) \leq Q_{min}\right\} \tag{9}$$

on conditions that

$$\begin{cases} \dfrac{\partial \upsilon(x,t)}{\partial t}+\dfrac{\partial Q(x,t)}{\partial x}=q(x,t), \\ q(x,t)=-\sum\limits_{i=1}^{n} q_i \delta(x-a_i) l(t-T), \\ Q(x,0)=Q_0(x), \\ x\ge 0, \quad t\ge 0, \quad \nu>0, \end{cases} \tag{10}$$

control action

$$Q(0,t)=Q_1(t). \tag{11}$$

For lateral water intakes, the quality factor of the water distribution process can be used to choose the root-mean-square integral deviation of the actual water flow from the planned (limited) values for the water distribution period $[0,T]$

$$I_1=\sum_{j=1}^{N}\int_0^T \left(q_j(t)-q_j^*\right)^2 dt \to min,$$

where, $q_j(t)$ – the actual value of the water flow of the j th lateral water intake; q_j^* -the planned value of the water flow of the j th water intake.

The minimization task (9) on condition (10) using the control action (11) presents oneself a task of controlling systems with distributed parameters [5].

Substituting the solution $Q(x,y)$ of the Green function defined by (6) into (9), we obtain [9]

$$I=\int_0^T\int_0^l\left[\int_{t_0}^t\int_0^l G(x,\xi,t-\tau)w(\xi,\tau)d\xi d\tau-Q^*\right]^2 dxdt \to min, \tag{12}$$
$$\Omega=\left\{R^N \mid Q_{min}\le Q^*(x,t)\le Q_{max}\right\}.$$

Further, substituting the standardizing function (6) in (12), we obtain [10]

$$\int_0^T\int_0^l\left[\int_{t_0}^t\int_0^l G(x,\xi,t-\tau)\left(q(\xi,\tau)+Q_0(\xi)\delta(\tau)+\right.\right.$$
$$\left.\left.+\nu Q_1(\tau)\delta(\xi)\right)\int_0^l d\xi d\tau-Q^*\right]dxdt \to min, \tag{13}$$
$$\Omega=\left\{R^N \mid Q_{min}\le Q(x,t)\le Q_{max}\right\}.$$

With a glance of the form of the Green's function, $q(x,y)$ and with $Q_0(x)=0$, considering the function $Q_1(t)$ in the form of a step function, the minimum of functional (13) is achieved when controlling

$$Q_1(t)=\sum_{i=1}^{n} q_{n-i} l\left(t-\sum_{k=1}^{i}\tau_{n-k}\right). \tag{14}$$

In this case, the solution of equation (15) is written in the analytical form

$$Q(x,t)=-\sum_{i=1}^{n} q_i l(x-a_i) l\left(\upsilon(t-T)-(x-a_i)\right)+$$
$$+l(t-\upsilon x)\sum_{i=1}^{n} q_{n-i} l\left(t-\sum_{k=1}^{i}\tau_{n-k}\right). \tag{16}$$

The physical meaning of this solution is to ensure changes in water flow in the lateral water intakes in different transits of the canal, it is necessary to change the water flow in advance at the beginning of the section.

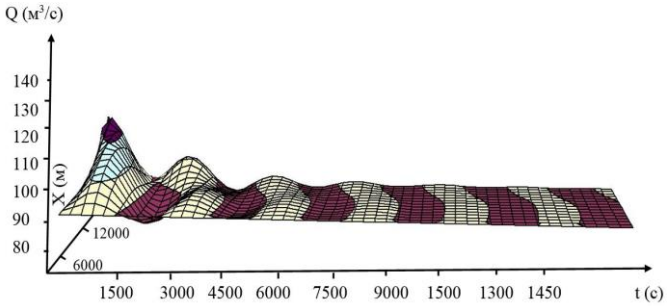


Fig. 3. Changing water consumption over time and length of the canal.

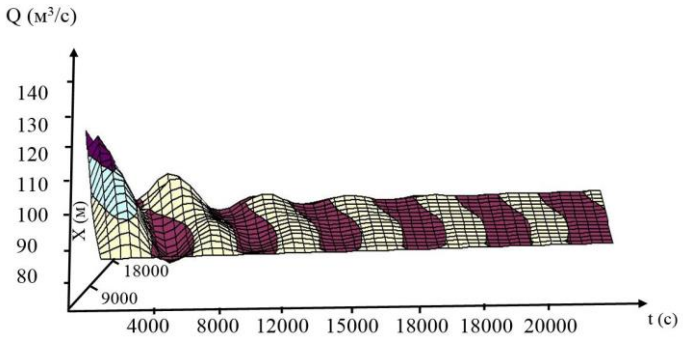


Fig. 4. Changing water consumption over time and length of the canal.

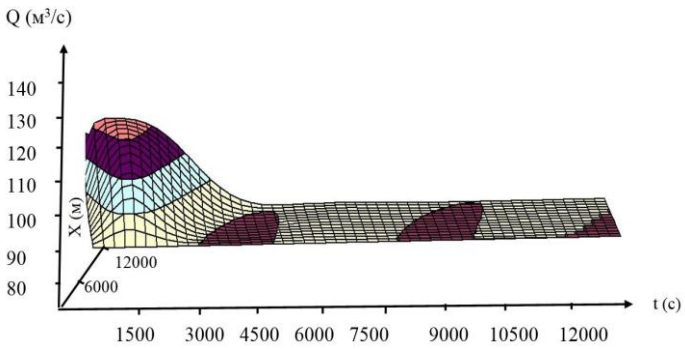


Fig. 5. Changing water consumption over time and length of the canal.

From the drawings it is clear that, after opening the shutters, the increased consumption at the beginning of the section of the channel allows to increase the water level along the length of the specified section of the channel.

3. CONCLUSION

Mathematical models of optimal water distribution in the main channels of irrigation systems (direct wave model, model of kinematic free unestablished water movement on the channel) with corresponding limitations and control effects.

The results of numerical experiments show that, after the opening of the shutters of the hydraulic structure at the beginning of the section of the canals, the increased consumption allows to increase the water level along the length of their sections. The levels and costs of water at the end of the canal sections eventually stabilize at the specified level, which is necessary for the water intakes from the canals located there. This confirms that the mathematical model developed by us is workable and can be used to optimally manage the distribution of water in the various self-flowing canals of the country's water industry.

REFERENCES

[1] Wang G.-T., Chen S. A., "Semianalytical solution of the Saint-Venant equations for channel flood routing", Water Resources Research, Vol.39, No.4, doi: 10.1029/2002WR001690, (**2003**), pp. 1-10.

[2] Valérie Dos Santos, Mickael Rodrigues, Mamadou Diagne, "A multi-models approach of Saint-Venant's equations: a stability study by LMI", International Journal of Applied Mathematics and Computer Science, De Gruyter, 22 (3), https://hal.archives-ouvertes.fr/hal-00701005. (**2012**), pp.539-550.

[3] Barros R. M., "Case studies for solving the Saint-Venant equations using the method of characteristics: pipeline hydraulic transients and discharge propagation", // IOP Conf. Ser.: Earth Environ. Sci, (**2019**), pp.1-10.

[4] Butkovsky A.G., "Characteristics of systems with distributed parameters", - M.: Science, (**1979**), 225p.

[5] Abutaliev F.B., Rakhimov Sh.Kh., Begimov I., "Optimal control of machine water reservoir systems", Tashkent: Fan, (**1992**), 152p.

[6] A.V. Kabulov, A.J. Seytov, A.A. Kudaybergenov. "Classification of mathematical models of unsteady water movement in the main canals of irrigation systems", International Journal of Advanced Research in Science, Engineering and Technology, Vol.7, Issue 4, (2020).

[7] Kabulov A V Normatov I H and Ashurov A O 2019 Computational methods of minimization of multiple functions Journal of Physics 1260

[8] Kabulov A V and Normatov I H 2019 About problems of decoding and searching for the maximum upper zero of discrete monotone functions Journal of Physics 1260

[9] Normatov I H 2018 Principle of independence of continuation of functions multivalued logic from coding Journal of Physics 1210

[10] Kabulov A V, Normatov I H and Karimov A 2020 Algorithmization control of complex systems based on functioning tables Journal of Physics 1441

An automated agricultural shading for crops with multiple controls

Jafrina Jabin
Department of Electrical & Computer Engineering
North South University
Dhaka 1229, Bangladesh
jafrina.jabin@northsouth.edu

Ahmed Masud Chowdhury
Department of Electrical & Computer Engineering
Indiana University – Purdue University Indianapolis
Indiana 46202, United States
amasudch@purdue.edu

Erteza Tawsif Efaz*
Department of Electrical & Electronic Engineering
Ahsanullah University of Science & Technology
Dhaka 1208, Bangladesh
ertezatawsif@gmail.com

Md Ehtesham Adnan
Department of Electrical & Computer Engineering
North South University
Dhaka 1229, Bangladesh
ehtesham.adnan@northsouth.edu

Mohammad Rezaul Islam
Department of Electrical & Computer Engineering
North South University
Dhaka 1229, Bangladesh
mohammad.rezaul@northsouth.edu

***Abstract*—Optimum ecological conditions are essential for ideal plant growth and improve harvest yields. Some crops grow better in low-intensity sunlight and suffer from sunburn damages due to excessive sunlight. As such, these types of harvests are required to be cultivated under shades. This paper describes the design of an automated agro shading incorporating FRDM-K64F MCU with several control options that will free the farmers from constantly surveilling the corps for proper sunlight intensity. In this approach, there are multiple input methods to change the state of the shade flaps, including physical and automatic control via a developed android app for smartphones. Additionally, full automation can be turned on through Bluetooth connection to regulate the cover automatically based on the magnitude of the incoming sunlight in any given area. Thus, the shading mechanism can provide an optimal lux environment. For improving the yield of agricultural plants with nutritious values, the presented module will revolutionize agricultural automation – offering a cost-effective multi-controlled shading approach to solve farming issues regarding photosynthesis and weather conditions.**

Keywords—stepper motor, embedded system, internet of things, control architecture, android app

I. Introduction

Automation has transformed industries with substantial economic growth through its pervasive application almost in every sector. Despite this ubiquity of its implementation, there are still some vital sectors where automation has not been adopted at its full potential yet, and one such area is agriculture - among the first occupations of humankind since the dawn of civilization [1]. For an improved yield of crops, appropriate environmental conditions are required: having an optimum supply of water, fertilizer, soil, pesticides, and most of all – sunlight [2]. Within the conventional cultivation, a farmer needs to observe the intensity of incidental sunlight upon plants, which is different for the respective crops depending on the corresponding regions [3]. Several crops flourish under sunlight with low-intensity, thus requires to be cultivated within shades while there are crops that thrive in high-intensity light. Besides, excessive light radiation can damage fruits and vegetables as a consequence of sunburn [4].

One way to mitigate the mentioned issue is to create a greenhouse – construction by a roof built from cheap and translucent elements, like glass, in which the crops needing controlled atmospheric situations are often spread a lot easier [5]. These structures vary in volume from modest huts to industrialized structures, whereas a small glasshouse is acknowledged as a cold frame [6]. The inside of a greenhouse is revealed to daylight, which is why it grows significantly hotter than the outside temperature, providing another solution for shielding its contents in cold weather [7]. However, these approaches do not address the shading aspects for crops that grow in lower light intensity considering the open environment as well. When shading is utilized properly, it reduces sunburn injuries and increases crop yield highly [4].

This research aimed to expand the control system of a stepper motor to emulate the motion of shade for greenhouses. The control signal was transmitted through wireless communication using an android application that can be installed on a smartphone. Another method of input was through the on-board buttons that are present on the microcontroller board of Next Experience (NXP) Semiconductors.

The primary objective was to integrate the following features to the project,

1) Bluetooth android application.
2) Voice direction.
3) Gesture control.
4) System automation.
5) Manually position correction.

Various control mechanisms were used to regulate the stepper motor. Additionally, achieving conflict-free multiple input methods was an intention, meaning the system should be capable of handling several input signals coming in together. For this, the system was designed in such a way that only the current input signal will be executed first, and any other frequency would get ignored until the current execution is complete.

II. Background Analysis

A suitable degree of shading is compelling to form the most effective use of shades by improving and choosing proper

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

technologies. Within the extant procedures, crops are shaded by administering a black net with various mesh sizes for managing corresponding light energy. Automatic shading envisions monitoring and regulating of the climatic parameters, which directly or indirectly supervise the plant germination and so their yield. One very prominent question that arises is why not provide a permanent shading for greenhouses. The answer lies within the incontrovertible fact that permanent shading will leave all the plants underexposed, thereby making a decreased yield along with feebler nutrient value [8].

Though very few [1-3], existing automation technologies for agricultural shading lack several features regarding smart controlling. In 2017, Faruk Bin Poyen presented the latest – that is a sensor-based automated shading using Arduino UNO microcontroller as the brain of the system to control the shade (open-close) depending on the declared luminosity of the sunlight; missing out a handful of user-friendly options as well as dimming the incidental light [9]. By taking into consideration the possible aspects of today's scientific appliances [6-7], our proposed model is comparatively economical to integrate and smoother to maintain, offering multi-controlling options within the scope of adjusting the states of shade allowing reflective sunlight.

There is a high potential for the improvement of structural and controlling applications for greenhouses due to the increasing demand for cultivating crops year-round, which on the other end, will be highly beneficial to increase self-employed industries cost-effectively [10].

III. Embedded Systems Implementation

The small-sized low-cost NXP freedom development kinetics board, FRDM-K64F is selected as the Micro-controller Unit (MCU) – having the same form factor as Arduino UNO Rev3 pin layout that can be used instead of the latter for prototyping purpose to design Internet of Things (IoT) [11-13] for controlling the shade in terms of agriculture to cultivate plants.

Bluetooth is radio communication technology, utilized for transferring data between settled and portable gadgets covering small distances employing short-wavelength Ultra High Frequency (UHF) communication streams within the manufacturing, experimental, and pharmaceutical wireless bands (2.402-2.480GHz); making individual domain networks. Here, the employed HC-06 is a class 2 slave Bluetooth module intended for clear radio continual connection [14], which is connected with the MCU through the Receiver (RX) and Transmitter (TX) pins of the board. The HC-06 Bluetooth module connections are outlined in Table I as follows,

The motor driver serves as an interface among the motors and controller circuitry – requiring a large volume of current,

TABLE I. HC-06 Bluetooth module connection with FRDM-K64F

FRDM-K64F	HC-06
PTC16/D0 (UART3-RX)	TX
PTC17/D1 (UART3-TX)	RX
GND	GND
5v	VCC

whereas the controller circuit works on low current signals. Thus, the purpose of the employed L298N motor driver is to take in a low-current controller signal then convert it toward a higher-current, which will operate a motor through switching [15], where the motor rotation depends on the way that gate voltage is provided to input terminals. The L298N motor driver connections are outlined in Table II as follows,

Stepper motor, referred to as step/stepping motor, is a brushless DC motor that distributes a complete circle within many similar runs, which cannot run at a great pace but have a high holding torque. The Mitsumi M42SP-6NPK stepper motor is used for its 2 phase, 7.5° step angle, and bipolar driving excitation method with a rating of 12-24V/400mA [16]. The M42SP-6NPK stepper motor connections are outlined in Table III.

Laser Diode (LD) is a semiconductor unit almost like Light-emitting Diode (LED) within which a diode-pumped straight with current that can produce lasing situations at the junction of the diode [17]. As such, a hand gesture equipment is built with 2 LDs (LD1 and LD2) and 2 Light Dependent Resistors (LDRs), namely LDR1 and LDR2, connected to the MCU. The hand gesture module connections are outlined in Table IV.

Photo-resistor – a passive component, decreases resistance about receiving luminosity on the sensitive surface of the component [18]. In this case, a daylight sensor is created using an LDR, namely LDR3, where the voltage read by AnalogIn of the board. The daylight sensor device connections are outlined in Table V.

A 9V rechargeable Nickel-metal Hydride (NiMH) battery utilized as the power supply, and 1kΩ resistors employed with

TABLE II. L298N motor driver connection with FRDM-K64F

FRDM-K64F	L298N
PTB9/D2	Input 1
PTA1/D3 (PWM)	Input 2
PTB23/D4	Input 3
PTA2/D5 (PWM)	Input 4
GND	GND
-	VCC to external 9V supply

TABLE III. M42SP-6NPK stepper motor connection with L298N

L298N	M42SP-6NPK
Output 1	Phase 1A
Output 2	Phase 1B
Output 3	Phase 2A
Output 4	Phase 2B

TABLE IV. Hand gesture module connection with FRDM-K64F

FRDM-K64F	Hand-gesture
5V	LDR1‖LDR2‖LD1‖LD2
PTB2/A0	LDR1 (GND)
PTB3/A1	LDR2 (GND)
GND	LDR1‖LDR2 (GNDs)

TABLE V. Daylight sensor device connection with FRDM-K64F

FRDM-K64F	Daylight-sensor
5V	LDR3
PTB10/A2	LDR3 (GND)
GND	LDR3 (GND)

the LDRs as a fixed reference point besides avoiding any potential short-circuit, followed by the connection to the common ground.

IV. Multiple Control Architectures

Multiple methods of regulation, including android application, voice direction, gesture control, system automation, and position correction, as shown in Fig. 1, were developed for the greenhouse shading.

A. Android Application

An Android application (app) was built using an accessible online tool – named MIT App Inventor (Fig. 2), integrated into the system as a User Interface (UI) to control the motor's angular position. Starting from UI design, to underlying programming,

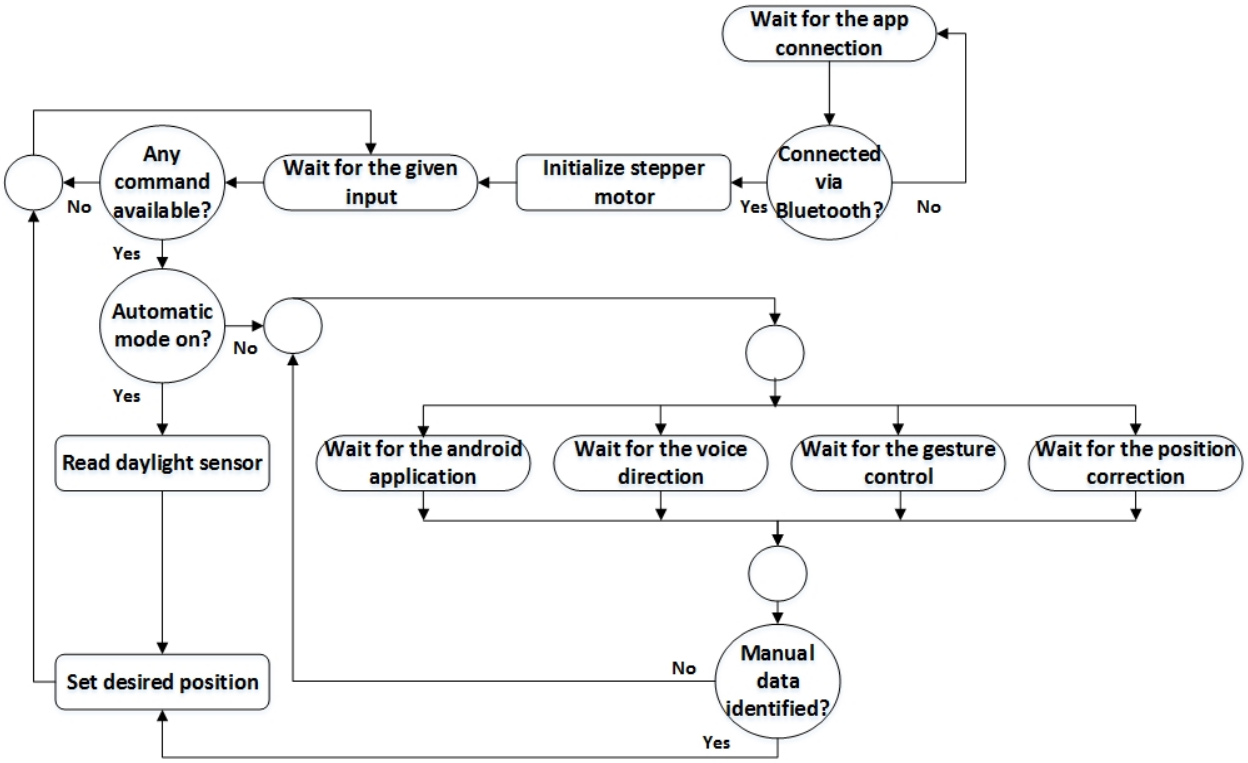


Fig. 1. Operational flow chart.

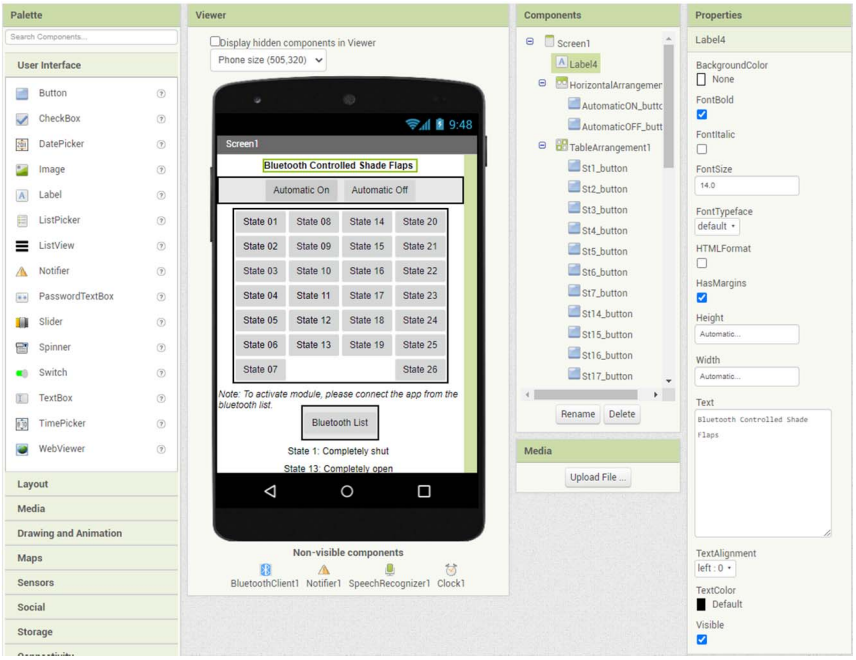


Fig. 2. MIT App Inventor designer interface.

all can be accomplished in a simple drag and drop manner by the development platform to publish the app in .apk format to be installed into any mobile device running on Android OS. An advantage of such an approach is that it allows for the rapid prototyping of IoT systems, which has proven to be perfect for the discussed control modules [19].

For prototyping the system, the Mbed Online Compiler was used that employs Mbed OS, providing an abstraction layer for microcontrollers and enabling its users with well-defined APIs compatible with C/C++ programming language to develop applications for any Mbed-enabled board. This compiler offers a UI, which provides a range of features, from code writing to compiling, and finally publishing it [20]. Once the code is written, the compiler returns a board-readable .bin file to be uploaded to the board through usual drag and drop.

The Tera Term was utilized to observe the motor state values, which is an open-source terminal software that permits users to establish serial communication between computer and microcontroller to send and receive data [21]. Moreover, the software was used to display whether the motor was rotating clockwise or counter-clockwise. Here, the app was developed in four parts (automatic, manual, connection process, voice command) (Fig. 3), where the button labeled 'Bluetooth List' establishes the connection between user's phone and the system, which must be confirmed first to enable the user to set the mode of operation either for manual request through the selection of any motor state (01-26) while the automatic function is turned off or for the automatic mode.

B. *Voice Direction*

To use the voice direction feature, the user needs to click on the 'Talk' button, as shown in Fig. 4, and speak out the state number that he/she wishes the motor to attain. The app recognizes this command and transmits the ASCII character required for the system to reach the position. Once the signal received by the Bluetooth module, it transfers the data through UART to the microcontroller board.

C. *Gesture Control*

For implementing the gesture control option, a separate module was built with a pair of LD as well as LDR to form the couples – set up adjacent to each other, as shown in Fig. 5, where the LDRs arranged in a voltage divider setting. The gesture control module continuously senses the laser beams for interruptions, and once the beam from the first couple is interrupted, the board awaits for the second beam to become obstructed for 500ms. If that happens, the program recognizes a sliding hand gesture towards the right, thereby executing a clockwise motion of the motor and vice-versa.

D. *System Automation*

In the case of system automation, the voltage divider rule was applied again, as shown in Fig. 6, indicating that the automation is achieved by sensing daylight data and comparing it to a threshold value to decide for emulating the opening/closing of the flaps. The user would have to click on the 'Automatic On' button of the app, which will stop the system from taking any other form of input – be it from the app or the board switches; where the automatic mode can be disabled by clicking the 'Automatic Off' button. When enabled, the flap opening motion is executed as LDR3 senses daylight; however, if it senses darkness, the flap closing motion executes.

In automatic mode, the system is programmed to rotate the shade flaps for four different states according to the intensity of the sunlight. As outlined in Table VI, the flaps can change the states (01, 05, 09, and 13) in respect with day to night (900 to 100lux). The resistance of an LDR decreases with increasing luminosity, leading the model to attain the corresponding state regarding the varied voltages.

E. *Position Correction*

While operational, the system might lose power suddenly, and when the power comes back on, the motor position that the system had attained before the power outage is set as the initial

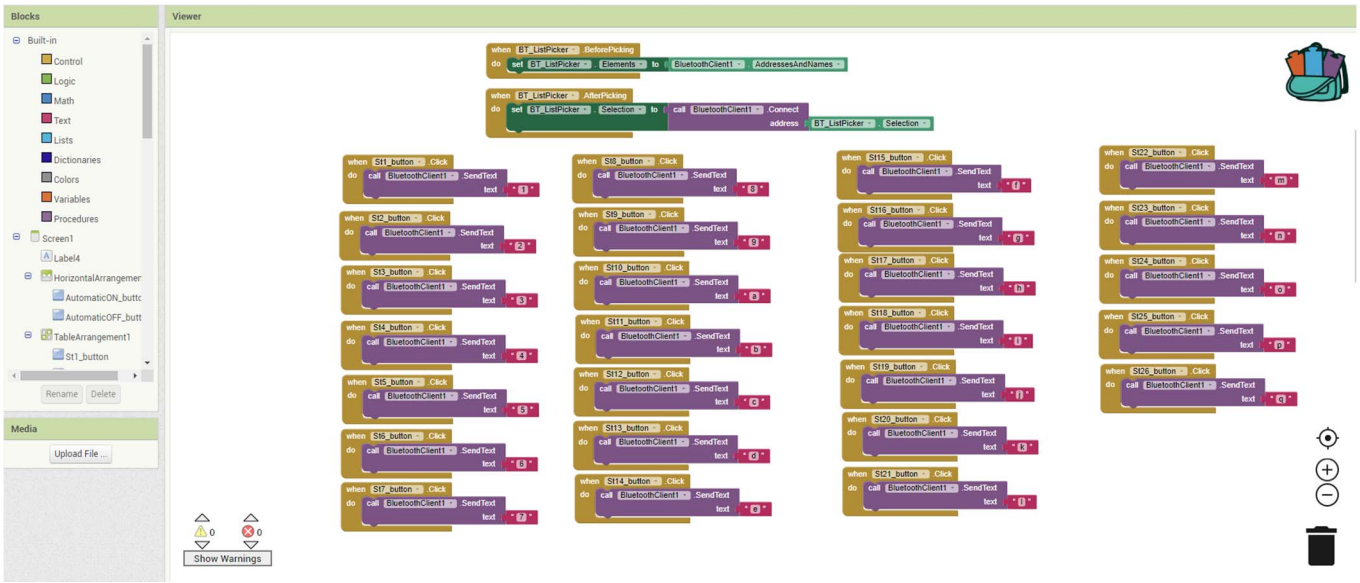


Fig. 3. MIT App Inventor blocks tab.

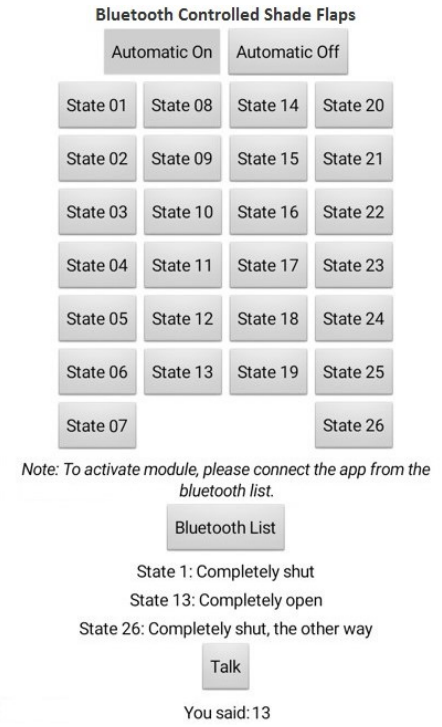


Fig. 4. Designed android app.

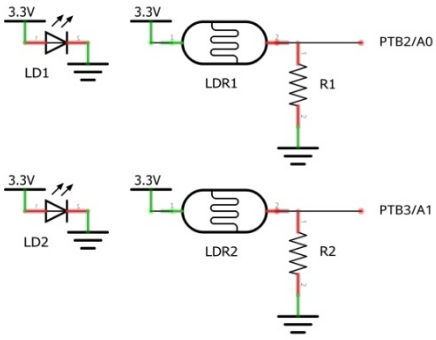


Fig. 5. Hand gesture module.

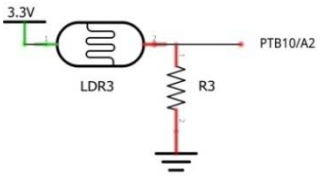


Fig. 6. Daylight sensor device.

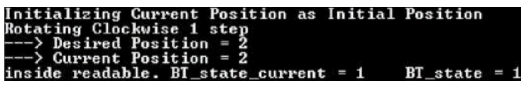


Fig. 7. Initial position correction.

TABLE VI. FLAP ROTATION IN AUTOMATIC MODE

CdS LDR resistance	State (rotation)	Condition of Shade
5kΩ at 900lux	01 (0°)	Fully closed
17.5kΩ at 650lux	05 (30°)	Partially open
32.5kΩ at 350lux	09 (60°)	Moderately open
45kΩ at 100lux	13 (90°)	Fully open

position – causing a loss of state synchronization. For correcting this, the system was equipped with a solution. Where the user can use the onboard switches to step towards the desired state. Once the motor gets to the desired state, the user can push the two switches simultaneously to set the attained position as the initial position (State 01). When the position correction is successful, as shown in Fig. 7, a message will be prompted on Tera Term stating, 'Initializing Current Position as Initial Position.'

Motor rotation can either be clockwise or counterclockwise or even both to function the opening and closing of the shade flaps – defined by the program, where 'State 01' implies 0° rotation of the shade flaps, 'State 13' implies 90°, 'State 26' implies 180°, and all other states imply the intermediate angles of rotation in terms of the stepper motor.

V. AGRICULTURAL SHADING

After defining a threshold value for the daylight sensor depending on the type of the plants and site – while setting the module, the system will be able to automate the opening (State 13) and closing (State 01 or 26) of the flaps of the shade. Besides, voice direction can become enabled by disabling the automated system, by which the amount of incidental sunlight can be controlled, referring to State 01-26, relying on various weather conditions. The gesture control mode can also become useful for emergencies as a means of immediate execution to close/open the cover from a distance and without even touching anything.

The estimated response time for different control systems is outlined in Table VII, showing the automatic mode with the minimum delay (0.24s). Though the voice direction mode requires the maximum time to respond (1.26s), it is one of the crucial modules for controlling the system due to its user-friendly compatibility. All the other modes of operation (android application, gesture control, and position correction) to control the shade through motor rotation responded under 1s.

The demonstrated module proved to be suitable for a range of practical applications – considering the difficulty of manual supervision, from the rooftop gardens in cities to farming lands in the villages, depicted in Fig. 8 (drawn in Fritzing) is the connection diagram of the circuit. Here, illustrated in Fig. 9 and 10 are the complete setup of the system and 3D model of a shade – designed in SOLIDWORKS; denoting the horizontal setting of the flaps (canceling carrying load) – capable to rotate in a fixed position by hook via the mounted embedded system on the top of any corner pillars that might require a generic gear for angular motion of the shade, depending on the application size. As the shading material, parts of semi-transparent

TABLE VII. REACTION TIME FOR DIFFERENT CONTROL ARCHITECTURES

Mode (control)	Duration (second)
Android Application	0.34
Voice Direction	1.26
Gesture Control	0.78
System Automation	0.24
Position Correction	0.84

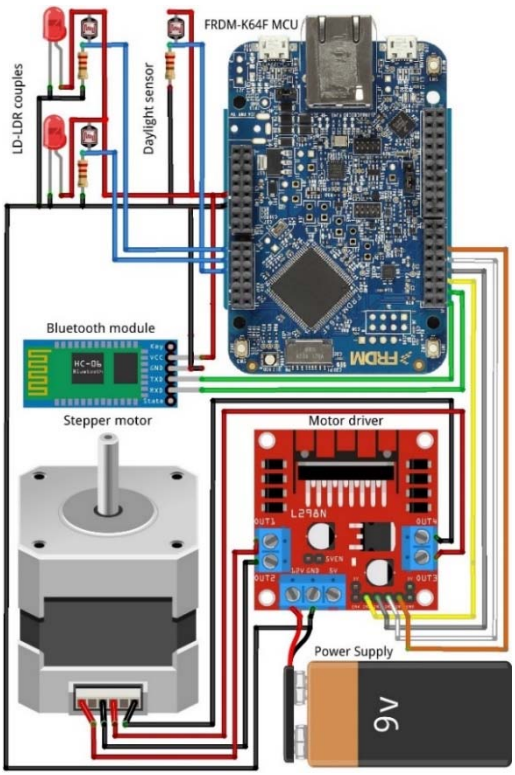


Fig. 8. Connection diagram of the circuit.

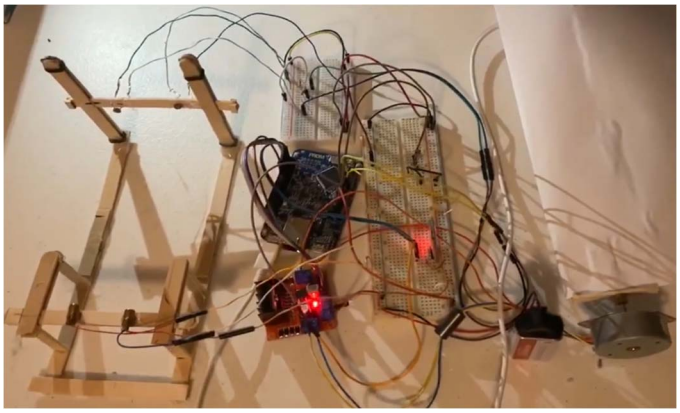


Fig. 9. Complete setup of the system.

polycarbonate sheets having 1mm thickness can be used considering proper photosynthesis in regards to the growth of the crops.

The input data and actuation method used in this technology can be applied in a plethora of scenarios. For example, stepper motor actuation is used frequently in controlling a robotic arm. The hand gesture module can be used in a setting where a touchless input is necessary; given the COVID-19 situation, a rail of gesture controlling couple could be utilized in elevators for floor selection to avoid touching high-risk surfaces.

Fig. 10. 3D model of a shade.

VI. Conclusion

In summary, throughout this, the possibility of controlling an automated agricultural plant shading system for better growth of crops to increase nutritious values along with productivity using the FRDM-K64F development board has been demonstrated. The first stage described the mechanism by which the control signals were sent through an android app to the stepper motor by Bluetooth to function the opening and closing of the shade as the motors rotated clockwise or counterclockwise. And, the second stage provided the necessary details about the improved version of the system by including a voice command and gesture control aspects to regulate the rotation of the motor. Here, the voice control is paramount in the design as the case of users living with some form of physical disabilities were taken into consideration. Through the integration of multiple hardware and software components in the developed project, it revealed how the embedded systems are shaping the modern world.

Finally, all the challenges concerning LD-LDR calibration and data conflict addressed to solve the problems. After debugging and troubleshooting, the actuation of the stepper motor to emulate the motion of the shade was achieved successfully. Once the model is implemented physically, the yield of crops could be determined. Furthermore, a comparative study could be conducted to record the yields under both conditions, i.e., with and without the shades. The dataset could further be used to train a neural network model by feeding in relevant parameters like daylight intensity, temperature, humidity, soil moisture, etc. to automate the shade flap, which in turn would facilitate a more comprehensive study.

ACKNOWLEDGEMENT

We would like to extend our humble gratitude towards Dr. Mohamed El-Sharkawy – Professor of Electrical & Computer Engineering, Purdue School of Engineering & Technology, for his resolute pedagogical support. Without his support, the compiled work would not have been possible. Furthermore, we also offer our sincerest appreciation towards the Internet of Things Collaboratory lab for all the hardware support which has proven paramount for all the testing and proof of concepts.

REFERENCES

[1] S. B. Zinnat and D. M. Abdullah, "Design of a fuzzy logic based automated shading and irrigation system," in *2014 17th International Conference on Computer and Information Technology, ICCIT 2014*, Mar. 2003, pp. 170–173, doi: 10.1109/ICCITechn.2014.7073098.

[2] A. Abdullah, S. Al Enazi, and I. Damaj, "AgriSys: A smart and ubiquitous controlled-environment agriculture system," in *2016 3rd MEC International Conference on Big Data and Smart City, ICBDSC 2016*, Apr. 2016, pp. 306–311, doi: 10.1109/ICBDSC.2016.7460386.

[3] W. Qiu, L. Dong, F. Wang, and H. Yan, "Design of intelligent greenhouse environment monitoring system based on ZigBee and embedded technology," in *Proceedings of 2014 IEEE International Conference on Consumer Electronics - China, ICCE-C 2014*, Feb. 2015, doi: 10.1109/ICCE-China.2014.7029857.

[4] T. Maughan, D. Drost, B. Black, and S. Day, "Using Shade for Fruit and Vegetable Production," *All Curr. Publ.*, Feb. 2017, Accessed: Jul. 16, 2020. [Online]. Available: https://digitalcommons.usu.edu/extension_curall/1654.

[5] "Greenhouse agriculture in Almería. A comprehensive techno-economic analysis - Publicaciones Cajamar." https://web.archive.org/web/20190209124321/http://www.publicacionescajamar.es/series-tematicas/economia/greenhouse-agriculture-in-almeria-a-comprehensive-techno-economic-analysis (accessed Jul. 04, 2020).

[6] "A brief history of cold frames | Cold Frame." https://web.archive.org/web/20100408002445/http://www.coldframe.org.uk/A-brief-history-of-cold-frames-19.aspx (accessed Jul. 04, 2020).

[7] "Shade house - See Harnois Greenhouse." https://web.archive.org/web/20160610204624/http://www.harnois.com/en/shadehouse/shade-houses.html (accessed Jul. 04, 2020).

[8] D. Or, "Stochastic Analysis of Soil Water Monitoring for Drip Irrigation Management in Heterogeneous Soils," *Soil Sci. Soc. Am. J.*, vol. 59, no. 5, pp. 1222–1233, Sep. 1995, doi: 10.2136/sssaj1995.03615995005900050003x.

[9] D. Or, "Stochastic Analysis of Soil Water Monitoring for Drip Irrigation Management in Heterogeneous Soils," *Soil Sci. Soc. Am. J.*, vol. 59, no. 5, pp. 1222–1233, Sep. 1995, doi: 10.2136/sssaj1995.03615995005900050003x.

[10] I. Ioslovich, P. Gutman, and I. Seginer, "A non-linear optimal greenhouse control problem with heating and ventilation," *Optim. Control Appl. Methods*, vol. 17, no. 3, pp. 157–169, Jul. 1996, doi: 10.1002/(SICI)1099-1514(199607/09)17:3<157::AID-OCA570>3.0.CO;2-X.

[11] "FRDM-K64F Platform|Freedom Development Board|Kinetis MCUs | NXP." https://www.nxp.com/design/development-boards/freedom-development-boards/mcu-boards/freedom-development-platform-for-kinetis-k64-k63-and-k24-mcus:FRDM-K64F (accessed Jul. 04, 2020).

[12] M. E. Adnan, N. M. Dastagir, J. Jabin, A. M. Chowdhury, and M. R. Islam, "A cost effective electronic braille for visually impaired individuals," in *5th IEEE Region 10 Humanitarian Technology Conference 2017, R10-HTC 2017*, Feb. 2018, vol. 2018-January, pp. 175–178, doi: 10.1109/R10-HTC.2017.8288932.

[13] J. Jabin, M. E. Adnan, S. S. Mahmud, A. M. Chowdhury, and M. R. Islam, "Low cost 3D printed prosthetic for congenital amputation using flex sensor," in *2019 5th International Conference on Advances in Electrical Engineering, ICAEE 2019*, Sep. 2019, pp. 821–825, doi: 10.1109/ICAEE48663.2019.8975415.

[14] R. Challoo, A. Oladeinde, N. Yilmazer, S. Ozcelik, and L. Challoo, "An overview and assessment of wireless technologies and coexistence of ZigBee, bluetooth and wi-fi devices," in *Procedia Computer Science*, Jan. 2012, vol. 12, pp. 386–391, doi: 10.1016/j.procs.2012.09.091.

[15] E. T. Efaz, A. Al Mamun, K. Salman, F. Kabir, S. N. Sakib, and I. Khan, "Design of an indicative featured and speed controlled obstacle avoiding robot," in *2019 International Conference on Sustainable Technologies for Industry 4.0, STI 2019*, Dec. 2019, doi: 10.1109/STI47673.2019.9068018.

[16] E. T. Efaz, M. M. Mowlee, J. Jabin, I. Khan, and M. R. Islam, "Modeling of a high-speed and cost-effective FPV quadcopter for surveillance," in *23rd International Conference on Computer & Information Technology, ICACIT 2020*, Dec. 2020.

[17] S. Nakamura, "The roles of structural imperfections in InGaN-based blue light-emitting diodes and laser diodes," *Science*, vol. 281, no. 5379. American Association for the Advancement of Science, pp. 956–961, Aug. 14, 1998, doi: 10.1126/science.281.5379.956.

[18] "How an LDR (Light Dependent Resistor) Works – Kitronik Ltd." https://kitronik.co.uk/blogs/resources/how-an-ldr-light-dependent-resistor-works (accessed Jul. 04, 2020).

[19] "MIT App Inventor | Explore MIT App Inventor." https://appinventor.mit.edu (accessed Jul. 13, 2020).

[20] "Introduction - Introduction to Mbed OS 6 | Mbed OS 6 Documentation." https://os.mbed.com/docs/mbed-os/v6.1/introduction/index.html (accessed Jul. 13, 2020).

[21] "Tera Term Help Index." https://ttssh2.osdn.jp/manual/4/en (accessed Jul. 13, 2020).

Development of an Economical SCADA System for Solar Water Pumping in Iran

Mohammad Zamanlou
M.Eng. Student
Faculty of Engineering and Applied Science
Memorial University
St. John's, Newfoundland, Canada
Email: mzamanlou@mun.ca

M. Tariq Iqbal
Professor
Faculty of Engineering and Applied Science
Memorial University
St. John's, Newfoundland, Canada
Email: tariq@mun.ca

Abstract – **In this paper, we developed a cost-effective SCADA system for a solar water pumping system in Iran. The SCADA is based on IoT versions and is comprised of a raspberry pi zero W, Arduino nano, camera, SIM 5320A 3G module, voltage, current, and light sensors. We used Node-RED to design a graphical user interface and published it securely to the worldwide internet. This allows a user to connect to the server via an IP address and monitor and control the system. The implementation of the project resulted in an open-source server and cost around CAD$ 162.38.**

Keywords— SCADA, IoT applications, Node-RED, Arduino, Raspberry pi, Wireless technologies

I. INTRODUCTION

Supervisory Control and Data Acquisition (SCADA) is a technology that enables electrical systems to be monitored and controlled from remote areas. SCADA technology has been around for many years, and various types of it were used in the industry to collect data from the system's electrical components such as batteries, electro-motors, etc. In recent years, the introduction and development of the Internet of Things (IoT) version of SCADA technology avail small to medium scale electrical projects such as solar water pumping of SCADA's advantages. It provides such projects with reliable and flexible control while being cost-effective to develop and install. [1-4]

In Iran, agriculture is facing many challenges due to the shortage of water and using new technologies such as drip-irrigation provides continued agricultural growth in the country. One of the applications of IoT based SCADA is in agriculture on solar water pumping systems. Such a SCADA system expands the development of a remote solar water pumping systems in the farming fields, such a system could be located at a far distance from living areas enabling farmers to efficiently operate the irrigation of their garden remotely.

Fig. 1 shows a schematic of a solar water pumping system. In this project, a SCADA system was developed to take measurements of various sensors such as PV voltage and current and control various switches for the submersible pump and photovoltaic panels. In addition, the SCADA provides monitoring for the environmental parameters such as a picture of the field and solar irradiance.

II. LITERATURE REVIEW

In [2], the authors presented a low-cost open source SCADA system to monitor and control a solar system where a customer can log-in to the dashboard and apply controlling over their solar system. They used a raspberry pi, an ESP32, a WiFi router and a few sensors. Also, they used thinger.io as their design hub. This system is only implementable in places where a WiFi is available also they are using a paid IoT hub named Thinger.io which adds up to the total cost of the system. We are advancing their system by adding a cellular network modem to make it possible to implement such system in remote areas and the designed project is programmed and secured completely on the local server using Node-RED. In [3], the authors designed a home automation system which shares a lot of similarities with the common IoT SCADA system. In their work, they made use of a raspberry pi, Arduino uno, and relays to implement controlling and monitoring through a dashboard. While this proposed system was successfully able to control the IoT devices, it uses Dataplicity to secure and expose the server to the internet which is a third-party application and puts a high risk over the server that data might be accessed through third-parties. Besides, their proposed system is not able to connect through a cellular network and is only available in urban areas. In our system, we

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

are using NGINX to secure and proxy pass the IP requests to the server and a 3G modem to connect to the cellular network [2, 3].

III. SCADA DESIGN

Fig. 2 shows a diagram of the proposed SCADA system. In this diagram, an Arduino nano is connected through a USB port to the raspberry pi and allows it to control sensors and LEDs connected to the ports of the Arduino. Also, a camera is used to picture the crop field or pumping system so that the user can monitor them. For wireless connectivity, we used both WiFi and 3G cellular network. The design procedure and explanation of the project are included in the following.

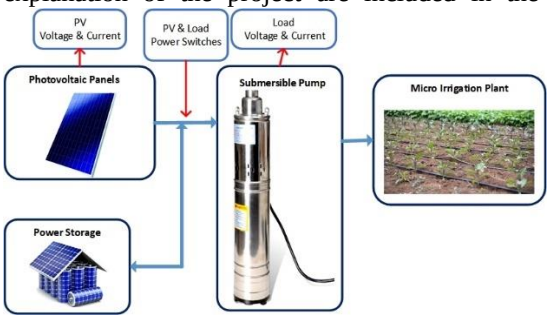


Fig. 1 – Schematic diagram of a solar water pumping system

A. WIRELESS CONNECTIVITY

A wireless connection is a great advantage for controlling systems in the era of modern technologies. In this regard, firstly, we took advantage of the raspberry pi's onboard WiFi module to connect it to the home router. Ref. 5 provides instructions to operate the raspberry pi without any need to plug in a monitor, mouse, or keyboard, which, in turn, made the system more cost-efficient. Secondly, we set up a cellular connection using instructions in [6, 7] to connect the server to a cellular 3G network. The cellular connection provides the server with the advantages of a cellular network, such as its availability.

B. WEB-BASED DASHBOARD

In our system, we used Node-RED in programing the raspberry pi to create a server. Node-RED is a programming environment using various types of nodes (i.e., serial port node) and mouse driven wiring of nodes. Arduino is controlled through Node-RED through a very user-friendly editor based on the internet browser. We used the Node-red-dashboard node to create a graphical user interface where the operator of the server can execute the system for monitoring and controlling purposes.

The final front end of our system is a web-based graphical user interface, so-called dashboard shown in figure 3. There are two tabs in the dashboard: Main tab and Environment tab; the main tab contains historical charts for presenting electrical data such as voltage and current, gauges for showing instant electrical measurements of voltage and current, and switches to control two LEDs which represent the power switches in the system. The environment tab includes a jpg picture box with a button to take a photo that allows the user to remotely monitor the field and sunlight data. Using this dashboard, a user has complete remote control over the system from anywhere in the world through 3G internet. Fig. 3A, and Fig. 3B shows the results of the dashboard that the user can interact with. [8-9]

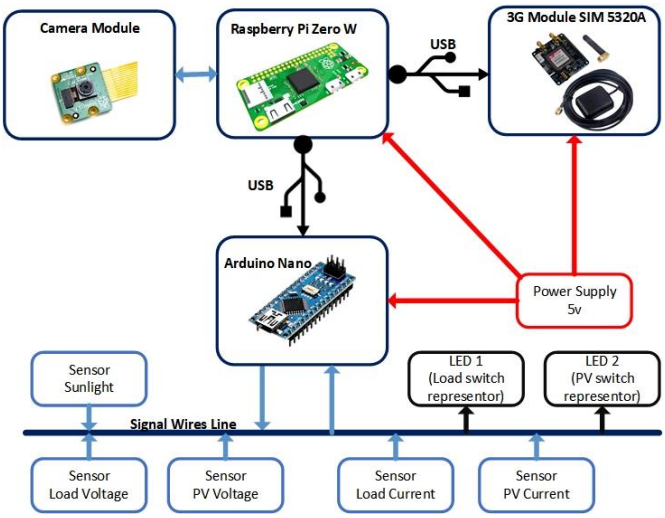


Fig. 2 – Block diagram of the server with various components

C. COMMUNICATION WITH ARDUINO NANO AND ELECTRICAL SETUP

Arduino nano is a small microcontroller board based on ATmega328 and has a very low power consumption. Using Arduino nano allows the server to sense and control various types of signals, such as analog signals for voltage measurement. There are many ways to communicate with an Arduino using Node-RED, such as Serial, Firmata, and Johnny-Five. We chose Firmata because it provides simplicity to the server and allows the computer to have direct access to the input and output pins of the Arduino.

In this setup, we designed a circuit to control two LEDs which represent load and PV power switches and sensors to measure load and PV voltages, load and PV currents, and sunlight.

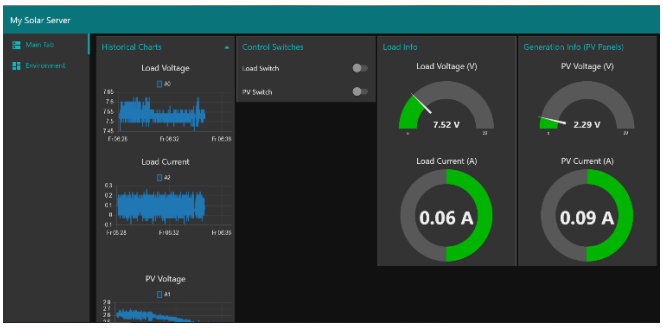


Fig. 3A - Dashboard Main Tab

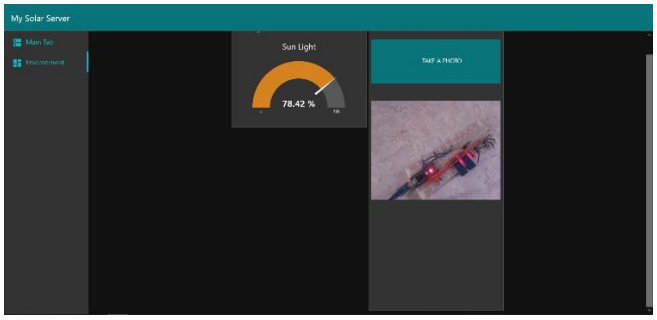


Fig. 3B - Dashboard Environment Tab

D. Exposing the Server to the WEB and Security

Since the user needs to remotely access the server, it needs to be published on a WEB server. On the raspberry pi, the NGINX web server application was used to create a web server and publish the dashboard on that server. However, publishing a local server on the world wide web where a tremendous number of threats exist, can put a very high risk on the server which is designed to monitor and control the solar water pumping system. In order to solve security concerns, we set up an HTTP authentication with NGINX; hence, the user needs to enter the username and the password of the server whenever connecting to the server IP address. All in all, using this method, the server is securely accessible through worldwide WEB and server IP is also known to the user. [10]

E. Cost Analysis

The designed system is very cheap to install and use. The cost of the system includes a low power raspberry pi zero w (CDN$ 37.60), an Arduino nano (CDN$ 4.80), and a camera (CDN$ 13.99), and a 3G internet module SIM 5320A (CDN$ 105.99). We should note that the raspberry pi zero W has an onboard WiFi module that allows it to connect to an internet router when available. In this case, we do not need a 3G module, and its price would be waived.

In addition, the cost of high-speed internet service is very negligible; for instance, the cost of a one-year plan with 48 GB internet is about CAD$ 6.5 from a carrier named Irancell in Iran. This advantage allows our system to be very profitable in the long run. The total cost of the system is between CAD$ 56.39 to CAD$ 162.38, depending on which internet type we choose to use.

There are many IoT hubs in the market that share similarities with our proposed system. As an instance, Microsoft Azure IoT Hub is a product that provides a platform to connect, monitor and control IoT devices. While the components required to implement an IoT SCADA application with Azure is almost at the same cost, it is required to pay an expensive monthly fee costing minimum 12.80 $CAD per each IoT Hub unit to implement the SCADA with it. It is worth mentioning that the number of messages communicated through IoT Hub in each plan is highly limited. This example shows the advantage of our designed system over other systems. Hence, the system is cost-effective in comparison with available IoT hubs such as Microsoft Azure in which the server is based on an internet cloud and are required to purchase their expensive plans. [11]

The experimental setup of the SCADA server is shown in Fig. 4.

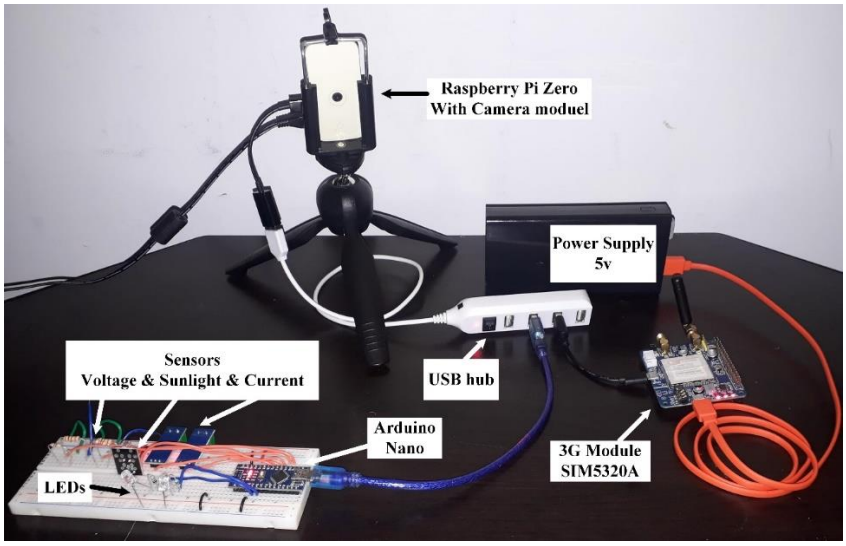


Fig. 4 - Experimental Setup of the SCADA server

IV. CONCLUSION

In this paper, the design of a cost-effective SCADA system for a solar water pumping system was presented. It consisted of a raspberry pi zero, Arduino nano, camera module, 3G module, and sensors for voltage, current, and sunlight and two LEDs. This server provides the system with a secure and complete monitoring and controlling GUI hub where the user can connect from anywhere through the internet and monitor and control a remote solar water pumping system. It is worth mentioning that the proposed system is implementable in any solar pumping system in the world where there is a connectivity to the internet available. The only difference will be the cost of a internet plan from the internet service providers.

Acknowledgment

The authors would like to thank Avrin Goom Azaran for providing funding for the research.

REFERENCES

[1] S. Phuyal, R. Bista, J. Izykowski, and D. Bista, "Performance Analysis of New SCADA Interface Developed in C# Environment," 2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 2020, pp. 1-4, doi: 10.1109/SCEECS48394.2020.209.

[2] L. O. Aghenta and M. T. Iqbal, "Development of an IoT Based Open Source SCADA System for PV System Monitoring," 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), Edmonton, AB, Canada, 2019, pp. 1-4, doi: 10.1109/CCECE.2019.8861827.

[3] C. Bepery, S. Baral and A. &. H. F. .. Khashkel, " Advanced Home Automation System Using Raspberry-Pi And Arduino," International Journal Of Computer Science And Engineering, vol. 8(2), pp. 1-10, 2019.

[4] S. R. Nath and R. Mandal, "Some studies on performance analysis of two different laboratory scale solar photovoltaic water pumping in irrigation systems," 2016 21st Century Energy Needs - Materials, Systems and Applications (ICTFCEN), Kharagpur, 2016, pp. 1-4, doi: 10.1109/ICTFCEN.2016.8052739.

[5] "How To Setup Raspberry Pi Zero W Headless WiFi" https://core-electronics.com.au/tutorials/raspberry-pi-zerow-headless-wifi-setup.html

[6] "How to setup a USB 3G Modem using usb_modeswitch and wvdial" https://tchekbo.wordpress.com/2018/03/31/how-to-setup-a-usb-3g-modem-using-usb_modeswitch-and-wvdial/

[7]" Raspberry Pi - Using a GPRS Modem for Network" https://www.youtube.com/watch?v=T2NBpfgfPaI

[8] "Node-RED", https://nodered.org/

[9] "Node red Dashboard" https://flows.nodered.org/node/node-red-dashboard

[10] "How To Set Up HTTP Authentication With Nginx On Ubuntu" https://www.digitalocean.com/community/tutorials/how-to-set-up-http-authentication-with-nginx-on-ubuntu-12-10#prerequisites

[11] "Azure IoT Hub pricing" https://azure.microsoft.com/en-us/pricing/details/iot-hub/

A C Code Generator for Fast Inference and Simple Deployment of Convolutional Neural Networks on Resource Constrained Systems

Oliver Urbann
Automation and Embedded Systems
Fraunhofer IML
Dortmund, Germany
oliver.urbann@iml.fraunhofer.de

Simon Camphausen
Automation and Embedded Systems
Fraunhofer IML
Dortmund, Germany
simon.camphausen@iml.fraunhofer.de

Arne Moos
Robotics Research Institute
TU Dortmund University
Dortmund, Germany
arne.moos@tu-dortmund.de

Ingmar Schwarz
Robotics Research Institute
TU Dortmund University
Dortmund, Germany
ingmar.schwarz@tu-dortmund.de

Sören Kerner
Automation and Embedded Systems
Fraunhofer IML
Dortmund, Germany
soeren.kerner@iml.fraunhofer.de

Maximilian Otten
Software and Information Engineering
Fraunhofer IML
Dortmund, Germany
maximilian.otten@iml.fraunhofer.de

***Abstract*—Inference of Convolutional Neural Networks in time critical applications usually requires a GPU. In robotics or embedded devices these are often not available due to energy, space and cost constraints. Furthermore, installation of a deep learning framework or even a native compiler on the target platform is not possible. This paper presents a neural network code generator (NNCG) that generates from a trained CNN a plain ANSI C code file that encapsulates the inference in single a function. It can easily be included in existing projects and due to lack of dependencies, cross compilation is usually possible. Additionally, the code generation is optimized based on the known trained CNN and target platform following four design principles. The system is evaluated utilizing small CNN designed for this application. Compared to TensorFlow XLA and Glow speed-ups of up to 11.81 can be shown and even GPUs are outperformed regarding latency.**

I. INTRODUCTION

A. Motivation and Related Work

Recent scientific advancements have led to a general acceptance of various classes of deep learning architectures as state of the art in machine learning, e.g. Convolutional Neural Networks (CNN). Research focused previously on image classification with a large number of classes [11] and is currently shifting towards object detection with approaches like R-CNN [4], where from the image region proposals are extracted and afterwards computed and classified utilizing CNN and SVM respectively.

CNN are thus a core component of a wide area of computer vision algorithms and are computationally expensive, usually accelerated by GPUs or FPGAs. However, recent advancements in mobile autonomous robotics as well as the Internet of Things (IoT) has opened a wide area of highly promising applications for these kind of algorithms, in which GPUs are not available and optimization of algorithmic performance becomes essential to save energy. At the same time those practical applications often come with the consequence that a large number of classes is not required and thus small CNN architectures are sufficient.

The first goal of this work is therefore to speed up the inference of small pretrained networks, e.g. for counting [15]. In general the motivation of this is twofold. First, the usual meaning is a high throughput given a large set of images to be classified. Overhead due to e.g. initialization is negligible. Second, and this is more important in this scope of application, the reduction of execution time - which correlates with latency as well as energy consumption. In mobile robotic applications latency is important for near-real-time reaction to sudden changes in the environment. In this case the latency should be as low as possible where computational time also correlates with energy savings. While the energy consumption is also a factor in mobile robotic applications its way more important in IoT applications, where it single handedly defines the lifetime of such a device. The set of images that must be classified at one time is rather low and thus the throughput is an suboptimal criteria for performance in this case. OpenVINO by Intel [5] is a good example for a highly optimized framework that addresses latency. They apply ahead-of-time (AOT) compilation and just-in-time (JIT) compilation for acceleration. For AOT compilation, they included the previously separately available nGraph compiler [3]. However, OpenVINO is limited to Intel platforms. With similar limitations, CompiledNN [14] also applies JIT compilation for Intel hardware to reduce the

This research has been funded by the Federal Ministry of Education and Research of Germany as part of the competence center for machine learning ML2R (01IS18038B).

latency.

As this paper shows, existing popular compiler like LLVM/-Clang and GCC perform poorly on finding parallelism to apply SIMD instructions. The goal of the Multi-Level Intermediate Representation (MLIR) [3] is to replace the LLVM IR to overcome this and other limitations. The Intermediate Representation Execution Environment[1] (IREE) is a first approach to apply MLIR for compiling machine learning models.

The second aspect is target platform deployability. Typically the network is embedded in a framework that provides images and processes the results. Embedding machine learning frameworks like TensorFlow [1] or Caffe [6] requires much overhead for an inference of a pretrained network. As a consequence, tools for generating object code for inference were developed for those kind of frameworks where TensorFlow XLA[2] and Glow[3] [10] are currently state-of-the-art. But their applicability to generic target platforms is limited. TensorFlow XLA generates object code that depends on TensorFlow code limiting the cross compilation capabilities for target platforms, whereas Glow's capabilities are currently limited to x86-64 and ARM64. It does not offer switches for other platforms, e.g. 32 bit targets, out of the box. TVM [2] is more flexible but also relies on LLVM limiting the possible platforms.

B. Contribution

In this paper we propose a neural network code generator[4] (NNCG) that generates C code from a trained CNN model. It focuses on the two relevant goals motivated previously:

- Generic scope of applicability and cross compilation for various target platforms
- Generation of fast executables allowing CNN inference on resource constrained systems (small robots, embedded microcontrollers etc.) on a CPU only

Generic Deployment

In contrast to common approaches which compile library code (e.g. Eigen) for operations like matrix multiplication and generate object code, we propose to generate plain ANSI C code. Since specialized code for each atomic operation (e.g. multiplication or addition) is generated and weights are included as constants, no libraries or prewritten code are needed except math.h and libmath. If desired target architecture dependent enhancements (e.g. SIMD instructions like SSE) can be utilized as well. As a result, the code can easily be compiled using a cross compiler or natively compiled on any target platform.

Fast Executables

Utilizing a math library and compiling to object code relies on a good optimization performance of the compiler as well as on the efficiency of the library. However, both library and optimizer are developed for any generic mathematical case. Instead, we exploit our knowledge about CNN in general and especially for the specific trained model to generate the most optimal code. Additionally, we intentionally choose C as output to fully benefit from the optimization capabilities of the compiler by generating code that is easy to optimize.

We identified four design principles to achieve those ideas, which we will discuss in detail in the following Sec. II:

- Loop unrolling and caching
- Conditional moves instead of branching
- Constants wherever possible
- Identification of applicable data structures for SIMD instructions

Usually the compiler should be able to handle most of these topics by itself. However, as the compiler has no background information this frequently fails in the field.

It has to be noted that these design principles limit the application of NNCG to small networks, as loop unrolling and floats written in ASCII text lead to large code files. E.g. MobileNetV2 [12] would require approx. 4 MB only for printing all weights in ASCII which leads to C files difficult to handle for a compiler.

In the evaluation in Sec. III we will show the advantage of addressing these points in NNCG based on small CNN adequate for our purpose. We compare the performance of NNGC with both above mentioned tools (TF XLA and Glow), wherever possible on a PC as well as a mobile robotic platform. We are able to show speed-ups of factor 1.41 up to 11.81 depending on network size and platform. We also compare the latency of a system with and without GPU. We can show that with small networks and a small number of images to classify the latency of our executable is many times smaller.

C. Structure

The following Sec. II describes the conceptual details of the NNCG and its implementation. Afterwards the results of NNCG are compared to the current state-of-the-art on various target platforms, in Sec. III. This evaluation will focus on mobile robotics as an application area, since in general it offers a larger variety of pattforms with different computational performance levels. The final Sec. IV concludes the paper and provides an outlook to future research.

II. Neural Network Code Generator (NNCG)

In this section we first describe the design principles (see Sec. I-B) in detail and continue how the CNN layer are realized fulfilling these principles.

The basic concept is the generation of C code from a trained Keras[5] model during an exemplary classification of an image. We reimplemented various CNN layer (all layer required for a custom YOLO [9] net) with focus on simplicity. During the calculation of each layer C code is written for all atomic operations, e.g. multiplication, addition, max operator etc. including the involved values as constants.

[1]https://github.com/google/iree
[2]https://www.tensorflow.org/xla/
[3]https://facebook.ai/developers/tools/glow
[4]https://github.com/iml130/nncg/
[5]https://keras.io

A. Design Principles

1) Loop unrolling and caching: In general, a loop consists of code for checking if a condition is met to continue executing the loop and a branch that repeats the loop. This has (mainly) two disadvantages: (1) Code for condition checking and branching and (2) negative effects on the pipeline of the processor resulting in a pipeline filled with wrong instructions if the processor cannot predict the condition correctly.

To mitigate this a compiler can unroll loops meaning the body of the loop is executed multiple times and the condition check and branch is thus executed less often. However, for this to work efficiently the number of loop iterations must be known, or further code is required to met the exact number of iterations.

On the other hand, unrolling results in more instructions that must be loaded from RAM which also affects the efficiency of the CPU cache. If all loops are unrolled completely, all instruction are only executed once and thus caching cannot increase execution speed.

Thus, we organize loop unrolling in different levels so that it can be chosen depending on the cache architecture of the target platform and the structure of the CNN. At level 0 all loops are unrolled. Level 1 does not unroll the outer most loop and so forth.

2) Conditional Moves: A typical operation is to copy a value into a register under some condition. In higher programming languages this usually is realized by a conditionally executed code block with a copy. It is skipped if the condition is not met again resulting in the clearance of the pipeline.

Thus, common processors implement copy instructions that are always executed but only actually copy the data if the condition is met. In worst case the time for executing this instruction is lost which is usually faster than refilling the pipeline.

Modern compiler should be able to identify candidates for a conditional move. However, as NNCG knows the semantics it can help by using the ternary operator known in C.

3) Constants: In common frameworks a CNN model is loaded into RAM during run-time and weights are passed to the calculation. The inference then must access these arrays using some addressing scheme. This may lead to unnecessary overhead as we can write the known constants into the corresponding line.

4) SIMD Instructions: Single Instruction Multiple Data (SIMD) instructions perform the same operation on multiple values and can thus speed-up the inference significantly. Modern compiler support these instructions but must be able to identify possible parallel calculations. To do so, the structure of the network must be known at compile time.

During code generation the structure of the calculations (matrix multiplications etc.) and the dimensions of vectors and matrices are known. Thus, parallel structures can be identified and SIMD instructions generated.

But, SIMD instructions are platform dependent. Currently we support Intel's SSSE3 and a general architecture without platform dependent code. However, other platform specific optimizations, such as for ARM's Cortex-M [7], can be integrated into the code generator as well.

B. Layer

We focus our work on layers required to implement a small YOLO [9] net. The following layers are also sufficient for other small networks that are suitable for embedded systems.

1) Convolution: Convolutional layers are the most computational demanding layers and thus a focus of this work. We support zero-padding and strides. Possible activations functions are the softmax function and (leaky) ReLU which we describe later.

To support padding we set all values to zero that are out of bounds by defining

$$\hat{x}_{ijk} = \begin{cases} x_{ijk}, & \text{if } 1 \leq i \leq h_{in} \ \wedge \ 1 \leq j \leq w_{in} \\ 0, & \text{otherwise,} \end{cases} \tag{1}$$

where x_{ijk} is the input of the convolution layer as a scalar at (i, j) and channel k, h_{in} the height and w_{in} the width. Applying this definition our implementation of the convolution can be written as

$$\begin{aligned} y_{ijk} &= y_{ijk} + \sum_{n=1}^{h_k} \sum_{m=1}^{w_k} \sum_{o=1}^{c_{in}} \omega_{nmok} \hat{x}_{i+n-p_t, j+m-p_w, k} \\ i &\in \{1, 1+h_s, \ldots, h_{in} - h_k + p_h + 1\} \\ j &\in \{1, 1+w_s, \ldots, w_{in} - w_k + p_w + 1\} \\ k &\in \{1, \ldots, c_{out}\}, \end{aligned} \tag{2}$$

with y_{ijk} as the output at (i, j) in channel k, h_k, w_k height and width of the kernel, c_{in} and c_{out} the number of input/output channels, ω_{nmok} the kernel weight at (n, m) for output channel k and input channel o, p_h, p_w the height and width of the padding, h_s, w_s the step height and width and h_{in}, w_{in} the dimensions of the input.

We see in Eq. 2 the calculation of a single value requires three nested loops. Furthermore, to calculate all output values three additional nested loops are required. The implementation of the first design principle is thus a trade-off between loop unrolling and code size. For the reasons explained in Sec. II-A1 unrolling all loops infinitely is only adequate for small networks and thus we follow a configurable approach. Currently, we support unrolling all loops with possible exceptions for the first and second outer loop and no unrolling.

To further specialize our code for different channel and spatial dimensions, we created multiple code versions of the convolution with different tradeoffs between cache utilization and register preassure. For each layer we independently benchmark every code version and select the one with the best runtime performance.

Implementation of design principle 3 depends on unrolling. If no loop is unrolled we generate an array containing all weights as constants. If loops are unrolled, the constants can be written into the corresponding code line.

For design principle 4 we identified the output channels (loop over k in Eq. 2) as a proper dimension for SIMD instructions. As can be seen, this loop does not affect the three

inner loops and is thus simple to apply. For SSSE3 the number of channels (in Eq. 2 denoted by c_{out}) should be dividable by 4 such that the number of filters in convolutional layers should be a multiple of 4.

2) Max-Pooling: The max-pooling layer searches for the maximum of all values in a two-dimensional window,

$$\begin{aligned} y_{ijk} &= \max\left(x_{i\cdot h_s, j\cdot w_s, k}, \ldots, x_{i\cdot h_s+h_k, j\cdot w_s+w_k, k}\right) \\ c_{out} &= c_{in} \\ i &\in \{1, h_{in} - h_k + 1\} \\ j &\in \{1, w_{in} - w_k + 1\} \\ k &\in \{1, \ldots, c_{out}\}. \end{aligned} \tag{3}$$

This two-dimensional window requires (in a basic form) two nested loops with additional three outer loops for each value of the output feature maps. Comparable to the convolution layer, we support no unrolling and full unrolling with possible exceptions for the outermost and second outermost loop. Furthermore, SIMD instructions are applied over channels if the number of filters in the previous convolution layer is a multiple of 4 (SSSE3).

3) (Leaky) ReLU: The ReLU layer consists of only three nested loops. We apply the same rules for unrolling as for max-pooling. The implementation of ReLU is simply,

$$y_{ijk} = \max\left(x_{ijk}, 0\right). \tag{4}$$

A leaky ReLU layer can mathematically described as:

$$y_{ijk} = \begin{cases} x_{ijk}, & \text{if } 0 < x_{ijk} \\ \alpha \cdot x_{ijk}, & \text{otherwise} \end{cases}, \tag{5}$$

where α is a factor realizing the "leaky" feature. The implementation for SSSE3 is also a max function with additional code for α. For a general architecture we utilize the conditional operator of the C language to implement the second design principle supporting the compiler utilizing conditional moves.

4) Batch Normalization: Batch Normalization was introduced to improve the performance of CNNs, as well as to stabilize the training process. The layer consists of a learnable affine transformation of the input feature map,

$$y_{ijk} = \frac{x_{ijk} - \mu}{\sigma}. \tag{6}$$

The calculation can be incorporated into a preceeding convolutional layer by modifying the weights and bias as shown below,

$$\begin{aligned} \mathbf{bn}\left(\mathbf{conv}\left(x\right)\right) &= \frac{\sum_i x_i w_i - \mu}{\sigma} \\ &= \frac{\sum_i x_i w_i}{\sigma} - \frac{\mu}{\sigma} \\ &= \sum_i x_i \left(\frac{w_i}{\sigma}\right) - \left(\frac{\mu}{\sigma}\right). \end{aligned}$$

TABLE I
BALL CLASSIFIER CNN.

Layer	#	Size	Stride	Padding
Input	1	16x16		
Conv	8	5x5	2x2	same
ReLU				
Max-Pool		2x2	2x2	
Conv	12	3x3		valid
ReLU				
Conv	2	2x2		valid
Soft-Max				

TABLE II
PEDESTRIAN CLASSIFIER CNN.

Layer	#	Size	Factor	Padding
Input	1	18x36		
Conv	12	3x3		same
ReLU				
Max-Pool		2x2		
Conv	32	3x3		same
Leaky-ReLU			0.1	
Max-Pool		2x2		
Conv	64	3x3		same
Leaky-ReLU			0.1	
Max-Pool		2x2		
Dropout			0.3	
Conv	2	4x2		valid
Soft-Max				

III. EVALUATION

This section evaluates the main goals of the this work as mentioned in Sec. I-B: simple deployment and fast executables. We compare NNCG with both tools mentioned in the introduction that have comparable intentions: TensorFlow XLA and Glow in versions available in December 2018, 1.12 and c27b61c respectively.

Common robotic platforms are based on CPUs at different performance level. As an example for cognitive mobile robotic applications the Robocup Standard Plattform League has been chosen. Its robot Nao by SoftBank Robotics[6] is a typical example of a small and cheap mobile robot, which intentionally lack a GPU to save energy. But if energy consumption, heat dissipation and cost are relatively neglectable, also a GPU can be integrated. We thus include various target platforms in our evaluation. A desktop processor Intel i7 8650U with Ubuntu 14.04, an energy efficient platform Intel Atom J1900 with Ubuntu 14.04, the Nao V5 by SoftBank Robotics (Intel Atom Z530) with a custom 32 bit Linux and the NVIDIA GPU GTX 1050 in a mobile system.

We evaluate both goals by presenting exemplary scenarios in simple robotic example applications: a ball detector for robot soccer, a pedestrian detector and a robot detector, all inferred on the mentioned target platforms. The CNN utilized for these purposes are described in Tab. I, Tab. II and Tab. III, respectively. Our evaluation is based on custom CNN designed to be small enough to lead to acceptable sizes of the C code file

[6]https://www.softbankrobotics.com/emea/en/nao

TABLE III
ROBOT DETECTOR CNN.

Layer	#	Size	Factor	Padding
Input	3	80x60		
Conv	8	3x3		same
Batch Norm.				
Leaky ReLU			0.1	
Max-Pool		2x2		
Conv	12	3x3		same
Batch Norm.				
Leaky-ReLU			0.1	
Conv	8	3x3		same
Batch Norm.				
Leaky ReLU			0.1	
Max-Pool		2x2		
Conv	16	3x3		same
Batch Norm.				
Leaky ReLU			0.1	
Conv	20	3x3		same
Batch Norm.				
Leaky ReLU			0.1	

which is also desirable in terms of inference speed on mobile platforms. For example, a MobileNet V2 leads to an 78 MB C code file. We are still able to compile and run this file. However, we suggest smaller networks and thus we evaluate utilizing the networks presented. The CNN structure is chosen such that decent classification results can be achieved and the networks are adequate for a simple application on embedded devices.

To evaluate the first goal of this work, we give a subjective and comparative overview about simplicity and applicability of the tools to generate an executable of the mentioned pretrained networks (ball and pedestrian detector). Afterwards, the second goal is evaluated by comparing the time required to infer a single image on CPU and GPU using NNCG, TensorFlow XLA and Glow. Besides this we also show how single features of NNCG can lower the latency. We are also interested in how a GPU could perform if no overhead is present. We thus additionally evaluate the throughput of the GPU by applying a large set of images on the GPU and compare this speed per image with the tools on other platforms.

A. Training

For each scenario we train the CNN presented above utilizing realistic datasets.

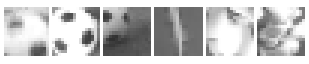


Fig. 1. Three positive examples (left) and three negative examples (right) of the ball dataset.

Ball CNN: The CNN presented in Tab. I is utilized in a pipeline for ball detection comparable to an R-CNN [4] in robot soccer. As a first step possible ball regions are extracted [13]. For this the image is first traversed along scanlines and segmented. On the resulting ball segments, multiple scanlines are created to find ball edge points. These in turn are used for circle fitting leading to a ball candidate for the presented CNN which is used for feature extraction and verification. An average of 20 ball candidates is created per image.

The size of the CNN can be very small for multiple reasons. A ball is an object with high contrast (white with black spots) and the appearance is invariant with respect to orientation.

The dataset consist of 455107 images with 125615 balls at a resolution of 16x16, see Fig. 1 for some examples. With 5% of the images for evaluation our trained CNN has an accuracy of 99.975%.

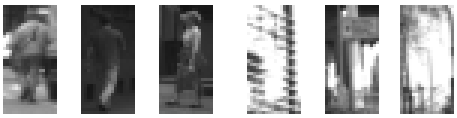


Fig. 2. Three positive examples (left) and three negative examples (right) of the Daimler pedestrian dataset.

Pedestrian CNN: In real world scenarios pedestrian or human detection is an important application. Humans are significantly harder to detect than a ball and we selected this as an example application to compare NNCG utilizing larger CNN, see Tab. II. For training we selected the Daimler pedestrian dataset [8], which consist of 49000 images with 24000 images of humans at a resolution of 18x36 per image, see Fig. 2 for example images. With 10% of the images for validation we achieve an accuracy of 99.02%.

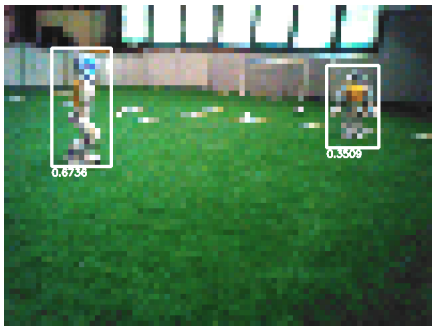


Fig. 3. Two robots (Nao by SoftBank Robotics) detected in a 80x60 pixel image of a soccer field.

Robot detector: We present above our ball detection application example, which is similar to the known R-CNN approach. Instead, for a robot detector application we build a pipeline based on the YOLO V2 approach [9] which is our third application example. In this paper we limit our presentation to the CNN utilized in the pipeline as described in Tab. III.

B. Generic Deployment

In this section we present different application scenarios and utilize NNCG, TensorFlow XLA and Glow to deploy executables including required steps to compile and link for the target platform. We study if the utilized tool is applicable under the circumstances of the scenario and show the simplicity by collecting the steps required for deployment.

Native Compilation for Host Platform: This is the most simple scenario in this evaluation as all tools are able to generate code and compile natively. Additionally, source code and libraries for compilation of the tools are also available. The host is an Ubuntu 18.04.1 LTS 64 Bit.

NNCG generates a C source code file that can be compiled to an object file. There are no dependencies for the robot detection CNN except for SSSE3 intrinsics on Intel platforms (`emmintrin.h`). The ball and robot classification additionally depends on `math.h` and `libmath` caused by exponentional functions in Softmax. Thus, all ANSI C compiler should be able to compile the C source file to an object file for a general architecture. Alternatively, if can be included in a project environment (CMake, Visual Studio etc.).

TensorFlow XLA includes the tool `tfcompile` to generate object files from a trained and stored CNN. It thus includes one more step than NNCG, the compilation of the code utilizing clang. However, the object file depends on many functions and the Eigen library shipped with TensorFlow. Thus, it is advisable to link this file to an executable within the TensorFlow environment providing all dependencies.

Glow's tool `image-classifier` generates an object file utilizing clang as well. It does not depend on libraries as TensorFlow XLA making the linking process as easy as NNCG on this platform. However, as the output is an object, compilation is limited to platforms supporting clang. Furthermore, Glow does not support all layer required for a CNN based on the YOLO approach, namely leaky ReLU.

Deployment on Atom (J1900) with similar OS: In this scenario the host platform for compilation is the same as above with a different target platform. Two limitations differ this scenario. First, the target CPU only supports a limited subset of SIMD instruction compared to the host (SSSE3). Second, Ubuntu is installed in Version 14.04.5 LTS.

The C code file generated by NNCG can be compiled natively on the target platform as it only requires a basic C compiler installation. Alternatively, it can be compiled on the host with static linkage and by specifying the target architecture (`bonnell` here).

TensorFlow XLA also supports the specification of a target platform. Static linkage is possible including the dependencies to TensorFlow and Eigen. However, a native compilation would require to install TensorFlow on the target platform and is thus not considered here.

Glow's `image-classifier` does not allow to specify a different target platform. Thus, the generated object file contains AVX commands as these are available on the host but not on the target platform resulting in not working executables. Installation of Glow on the target platform was not considered here.

Deployment on Atom (Z530) with different OS: This is the platform of the Nao robot with a preinstalled OS. The CPU is more limited but supports the same SIMD extensions as above. Main difference here is the custom Linux distribution with 32 bit kernel. It does not provide a compiler, thus native compilation is impossible.

C source generated by NNCG can be cross compiled by specifying a 32 bit target and static linkage. In contrast, the object generated by TensorFlow XLA depends on Eigen source that cannot be compiled for 32 bit targets. Glow suffers the same limitations as above and is not applicable here.

C. Fast Executables

The previous section demonstrates the applicability for different platforms. In this section we continue the evaluation by measuring the required time for the inference. We measure the time required to classify a single image (ball or pedestrian) and for detecting robots in an image. We ran small networks 100.000 times and larger networks like the robot detection 1000 times and use the mean value. For each application example the results can be seen in Tab. IV, V and VI. As described in the previous section, some approaches are not applicable and thus no time measurement is available.

As can be seen, the speed-up factor of NNCG compared to TensorFlow XLA is between 1.41 and 11.81, compared to Glow 3.29. This also confirms the results of [10]. Additionally, we evaluated the ball and pedestrian CNN on a GeForce GTX 1050 GPU by NVIDIA using an executable by TensorFlow XLA. As can be seen, the overhead to utilize a GPU is tremendous for small CNN and does not change significantly for under 100 images classified at once.

As described in Sec. II-A, two features of NNCG are configurable: the architecture dependent SIMD extensions and loop unrolling. We can therefore evaluate the acceleration due to these features by first using a general architecture without SIMD extensions where both outer loops are not unrolled. We do this for the ball classifier on the i7 platform. The compiler (clang 6.0.0) is nevertheless enabled to use SIMD extensions and perform loop unrolling. However, it can be seen in Tab. VII that the speed-up factor of applying SIMD instructions as described in Sec. II-A4 is 4.9. If NNCG unrolls all loops there is an additional speed-up of 26%. This shows that the compiler is not able to find the optimum automatically.

TABLE IV
EXECUTION TIME OF BALL CLASSIFIER.

Platform	NNCG	Glow	TensorFlow XLA
Intel i7 (8650U)	2.10µs	7.53µs	24.81µs
Intel Atom (J1900)	17.51µs	N/A	69.12µs
Intel Atom (Z530)	46.50µs	N/A	N/A
NVIDIA 1050	N/A	N/A	5630µs

TABLE V
EXECUTION TIME OF PEDESTRIAN CLASSIFIER.

Platform	NNCG	Glow	TensorFlow XLA
Intel i7 (8650U)	135.7µs	N/A	191.8µs
Intel Atom (J1900)	1020.3µs	N/A	1757.2µs
Intel Atom (Z530)	2938.6µs	N/A	N/A
NVIDIA 1050	N/A	N/A	5762µs

TABLE VI
EXECUTION TIME OF ROBOT DETECTOR.

Platform	NNCG	TensorFlow XLA
Intel i7 (8650U)	474µs	2457µs
Intel Atom (J1900)	1109µs	6797µs

TABLE VII
SPEED COMPARISON OF DIFFERENT FEATURES.

General	SSSE3	SSSE3 and Full Unroll
12.94µs	2.64µs	2.10µs

IV. CONCLUSION AND OUTLOOK

This paper presents a neural network code generator NNCG that writes ANSI C code for a trained CNN. We shown that embedding this file or a compiled object is a simple task and allows to deploy the CNN on all platforms that provide an ANSI C compiler or that can be a target platform of a cross compiler. Additionally, NNCG can exploit that the structure and gains of the CNN are know at generation time resulting in executables up to 11.81 times faster than previous well-known approaches.

Future work will focus GPU kernel code and more layer types to support modern widely known CNN structures. Furthermore, currently only SSSE3 is a supported SIMD instruction set. An extension of NNCG to other instruction sets like AVX or NEON can be realized rapidly.

REFERENCES

[1] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X.: TensorFlow: Large-scale machine learning on heterogeneous systems (2015). URL https://www.tensorflow.org/. Software available from tensorflow.org

[2] Chen, T., Moreau, T., Jiang, Z., Zheng, L., Yan, E., Shen, H., Cowan, M., Wang, L., Hu, Y., Ceze, L., et al.: {TVM}: An automated end-to-end optimizing compiler for deep learning. In: 13th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 18), pp. 578–594 (2018)

[3] Cyphers, S., Bansal, A.K., Bhiwandiwalla, A., Bobba, J., Brookhart, M., Chakraborty, A., Constable, W., Convey, C., Cook, L., Kanawi, O., et al.: Intel ngraph: An intermediate representation, compiler, and executor for deep learning. arXiv preprint arXiv:1801.08058 (2018)

[4] Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580–587 (2014)

[5] Gorbachev, Y., Fedorov, M., Slavutin, I., Tugarev, A., Fatekhov, M., Tarkan, Y.: Openvino deep learning workbench: Comprehensive analysis and tuning of neural networks inference. In: Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 0–0 (2019)

[6] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. In: Proceedings of the 22Nd ACM International Conference on Multimedia, MM '14, pp. 675–678. ACM, New York, NY, USA (2014). DOI 10.1145/2647868.2654889. URL http://doi.acm.org/10.1145/2647868.2654889

[7] Lai, L., Suda, N., Chandra, V.: Cmsis-nn: Efficient neural network kernels for arm cortex-m cpus. arXiv preprint arXiv:1801.06601 (2018)

[8] Munder, S., Gavrila, D.M.: An experimental study on pedestrian classification. IEEE transactions on pattern analysis and machine intelligence **28**(11), 1863–1868 (2006)

[9] Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017, pp. 6517–6525 (2017). DOI 10.1109/CVPR.2017.690. URL https://doi.org/10.1109/CVPR.2017.690

[10] Rotem, N., Fix, J., Abdulrasool, S., Deng, S., Dzhabarov, R., Hegeman, J., Levenstein, R., Maher, B., Nadathur, S., Olesen, J., et al.: Glow: Graph lowering compiler techniques for neural networks. arXiv preprint arXiv:1805.00907 (2018)

[11] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L.: ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV) **115**(3), 211–252 (2015). DOI 10.1007/s11263-015-0816-y

[12] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4510–4520 (2018)

[13] Schwarz, I., Hofmann, M., Urbann, O., Tasse, S.: A robust and calibration-free vision system for humanoid soccer robots. In: Proceedings RoboCup 2015 International Symposium. Hefei, China (2016)

[14] Thielke, F., Hasselbring, A.: A JIT compiler for neural network inference. In: S. Chalup, T. Niemueller, J. Suthakorn, M.A. Williams (eds.) RoboCup 2019: Robot World Cup XXIII, *Lecture Notes in Artificial Intelligence*, vol. 11531, pp. 448–456. Springer (2019)

[15] Urbann, O., Stenzel, J.: A convolutional neural network that self-contained counts. Journal of Image and Graphics **7**(4) (2019)

Optical and Electrical Characterization of n-MoS_2/p-Si Heterojunction Diode

Sayantika Chowdhury, P. Venkateswaran and Divya Somvanshi*
Department of Electronics and Tele-communication Engineering, Jadavpur University, Kolkata-700032
divya.somvanshi@jadavpuruniversity.in

ABSTRACT: This work reports the optical and electrical characteristics of n- Molybdenum disulfide (MoS_2)/p-Si based heterojunction diodes. MoS_2 quantum dots (QDs) based thin film synthesized on the p-Si substrate by facile-colloidal synthesis process followed by the spin coating. The optical properties of MoS_2 thin film have investigated by excitation dependent photoluminescence (PL) spectra and UV–Vis absorption spectra. Absorption spectra show only one peak in the near-UV region with a wavelength of < 300 nm which is typical exciton characteristics of MoS_2 QDs. The excitation dependent PL spectra show a redshift in peak wavelength with an increase in the excitation wavelength from 360 nm to 480 nm. The electrical characteristics are determined by room temperature current-voltage (I-V) and capacitance-voltage (C-V) characteristics. The value of barrier height (φ_B) and ideality factor (η) from I-V is calculated as 0.23 eV and 1.95 with a good current rectification ratio of 7480 at ±2 V.

Index terms- Molybdenum disulfide, Quantum dots, heterojunction diode, Optical properties

I. INTRODUCTION

Transition-metal dichalcogenides (TMDCs) semiconductors (ex: MoS_2, WS_2, etc.) are considered as promising materials for future optoelectronic applications [1-5] due to their excellent physical [6], electrical [7] and optical [6, 8] properties. Among all the TMDCs, MoS_2 is a promising material with a bandgap of ~1.29 eV (indirect) in the bulk whereas it becomes the direct bandgap of ~ 1.9 eV in monolayer form [6, 9]. To date, several researchers proposed various synthesis processes of MoS_2 quantum dots (QDs). Mukherjee *et. al*. [10] reported the colloidal synthesis of MoS_2 QDs on silicon platforms which exhibited excellent rectification behavior. In recent years, MoS_2 QDs/Si-based electronic devices have received significant attention due to its low-cost synthesis method, and integration compatibility and size-dependent tunable optical properties [2, 11-13].

Here, in this work, we have synthesized MoS_2 QDs based thin film on p-Si substrates through a two-step colloidal-synthesis process for multifunctional optoelectronics applications. The optical characteristics of MoS_2 thin film are determined by excitation dependent PL spectra and UV–Vis absorption spectra. The electrical characteristics of n-MoS_2/p-Si heterojunction diodes are analyzed by room temperature I-V and C-V characteristics. Finally, all the results are discussed in detail.

II. EXPERIMENTAL SECTION

A. Materials:

Ammonium tetrathiomolybdate ($(NH_4)_2MoS_4$), Oleic acid (OA), Oleylamine (OLA) were purchased from Sigma-Aldrich, 1-octadecene (ODE) was purchased from Alfa Aesar and Acetone, Chloroform, Isopropyl alcohol was purchased from GR, Merck. The de-ionized (DI) water (ρ = 18 MΩ-cm) obtained from a Milli-Q water plant. The boron-doped p-Si wafer (100) of thickness ~380 μm was purchased from the TED PELLA, INC. used as a substrate.

B. Synthesis of MoS_2 quantum dots (QDs):

In this work, MoS_2 quantum dots (QDs) were synthesized by a facile colloidal approach. In the typical process, a mixture of $(NH4)_2MoS_4$ (0.013 cc), Oleic acid (OA) (1 ml), Oleylamine (OLA) (3 ml), and 1-octadecene (ODE) (6 ml) heated up to 120°C for 2.5 hours with stirring. The homogeneous solution was heated to 250 °C slowly in a dry nitrogen atmosphere and upheld at this temperature for the next 2.5 hours. Further, the temperature of the solution was cool down at room temperature. The resulting MoS_2 QDs were precipitated in acetone, rinsed with 2-propanol. Then stir the solution to mix well and then re-dispersed in 10 ml of chloroform. The schematic cartoon of the complete synthesis process has shown in Fig. 1(a). Further, as-synthesized MoS_2 QDs coated on p-Si substrates by spin coating with 3000 RPM for 20 sec. Before spin coating, p-Si substrates were cleaned by a standard RCA process.

C. Al/Ti/n-MoS_2/p-Si/Ti/Al heterojunction diode:

The ohmic metal contacts deposited on both sides of the n-MoS_2/p-Si heterojunction diodes. The aluminium (Al)/ titanium (Ti) (~40/70 nm) metal deposited on the front side of the n-MoS_2 thin film using a shadow mass. The uniform Al/Ti (~ 40/80 nm) thin film deposited on the back-side of p-Si substrates using physical vapour deposition technique (from Hind High Vacuum, India) at a vacuum level of ~ 4×10^{-6} mbar. The diameter of the deposited metal dots is 2 mm. The schematic diagram of the n-MoS_2/p-Si heterojunction in the forward bias operation is shown in Fig. 1(b). We note that the polarity of the source should be inverted for reverse bias operation.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

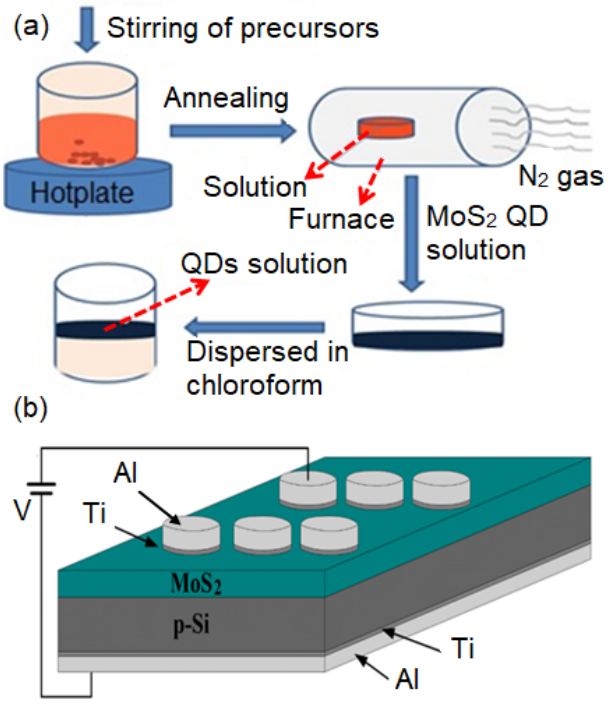


Fig. 1 (a) Schematic illustration of the colloidal synthesis process for MoS_2 Quantum Dots **(b)** Schematic diagram of Al/Ti/p-Si/n-MoS_2/Ti/Al hetrojunction diodes.

III. RESULTS AND DISCUSSION

A. Optical characteristics

The PL spectra (Omni PL, λ ~ 532 nm) of MoS_2 QDs are investigated under various excitation wavelengths as shown in Fig. 2 (a). It shows that, by increasing the excitation wavelength from 360 nm to 480 nm, a redshift in emission peak observed; the intensity also decreases rapidly with emission wavelengths varying from ~ 560 nm to 660 nm. The wavelength-dependent PL indicates polydispersity of the MoS_2 QDs distributions consistent with previous reports [14]. Another reason for this excitation dependent PL behavior is the defect level formation after oxygen adsorbed to the particle edges [15]. In multicolour imaging applications, this wavelength-dependent PL behavior is found very useful [16]. Fig. 2(b) shows the exciton wavelength versus the PL peak position, the peak position of PL increases with the increase of excitation wavelength. The linearly dependent PL behavior in MoS_2 QDs is also consistent with the previous results [14, 17]. As per literature, the size of QDs decreases with the increasing emission peak energy, in good agreement with the effect of quantum confinement [10].

The bandgap (E_g) energy of MoS_2 QDs based thin film has calculated as 3.88 eV as shown in Fig. 3. The absorption spectra of MoS_2 QDs based thin film coated on glass substrates is shown in the inset of Fig. 3. The absorption spectra of as-prepared MoS_2 QDs show only one peak in the near-UV region with a wavelength of < 300 nm which represents the typical exciton nature of MoS_2-QDs [18].

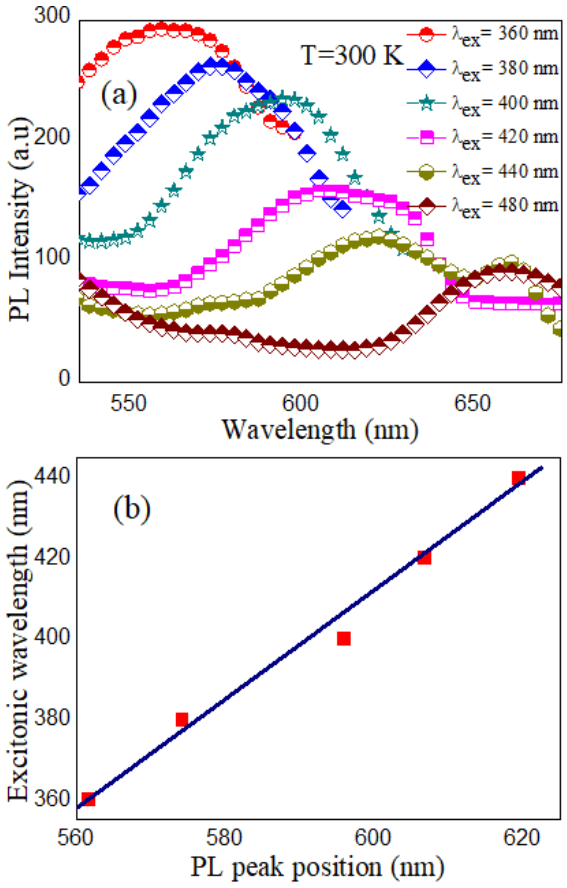


Fig. 2 (a) Excitation wavelength-dependent PL spectra of MoS_2 QDs based thin film on p-Si substrates. As excitation wavelength varies from 360 nm to 480 nm, a redshift in the main emission peak is observed, indicates polydispersity of the MoS_2 QDs **(b)** Monitored emission peak wavelength versus exciton wavelength.

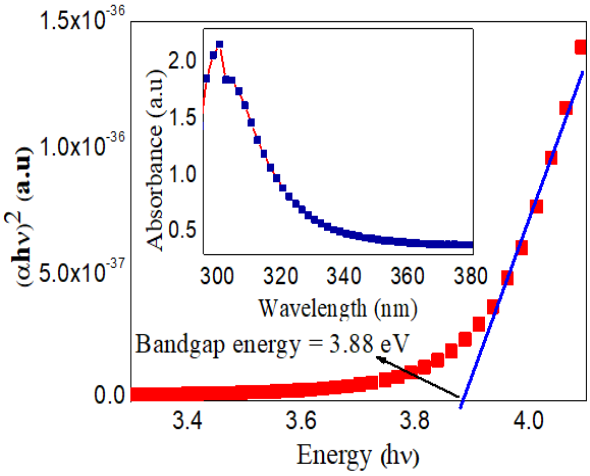


Fig. 3 Bandgap calculations of MoS_2 QDs coated on glass substrates; inset shows absorption spectra of as-prepared MoS_2 QDs coated on glass substrates.

B. Electrical Characteristics

We used a semiconductor parameter analyzer (Agilent B1500A) to analyze the I-V characteristics, in the voltage range from -2 V to +2V. The C–V characteristics of diodes measured at 1 MHz with LCR meter (HP-4284 A).

To understand the carrier transport across heterojunction, the energy band diagram of n-MoS_2/p-Si heterojunction diode in the thermal equilibrium condition is shown in Fig. 4. Based on the Anderson model [19], the value of conduction band offset (ΔE_c) and valence band offset (ΔE_v) has calculated which is given as ΔE_c = 0.15 eV, ΔE_v = 2.90 eV. Since $\Delta E_v > \Delta E_c$, thus the current transport in the n-MoS_2/p-Si heterojunction diode determined mostly by the flow of electrons from the n-MoS_2 to the p-Si side of the heterojunction diode.

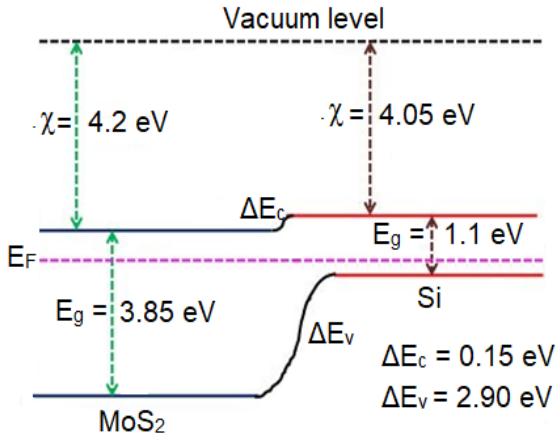


Fig. 4 Energy band diagram of n-MoS_2/p-Si heterojunction diode in thermal equilibrium condition

1) Current-Voltage (I-V) characteristics:

The I-V characteristics of the n-MoS_2/p-Si heterojunction diodes shown in Fig. 5. By assuming a pure thermionic emission model, the forward current at the n-MoS_2/p-Si heterojunction diodes can be expressed as [19]

$$I = I_0\left\{exp\left(\frac{q(V-IRs)}{\eta kT}\right) - 1\right\} \quad (1)$$

where q= *1.6×10^{-19}* is the electronic charge , V is the bias voltage, $\eta \sim (q/kT)\{dV/d(\ln I)\}$ is the ideality factor, k *(~1.38×10^{-23} J/K)* is the Boltzmann constant, R_s is the series resistance, T= *300 K*, and I_0 is the reverse saturation current defined as $I_0 = AA^*T^2\, exp(-q\Phi_B/kT)$ with A^* is the Richardson constant of MoS_2 ~ 54 A.cm^{-2}.K^{-2}, A is the contact area given as ~3.14×10^{-2} cm^2, and $\Phi_{B,IV} = (kT/q)\ln(AA^*T^2/I_0)$ is the effective barrier height (in volt) at zero bias. As-fabricated heterojunction diode shows a good rectification ratio (I_{on}/I_{off}) of 7480 at ± 2 V.

Fig. 5 shows the plot of ln (I)-V for determination of I_0, $\Phi_{B,IV}$ and η. The value I_0 can be calculated from the extrapolated intercept of ln I *vs.* V plot with the current axis $V = 0$. The value of $\Phi_{B,IV}$ and η from Fig. 5 are estimated as 0.23 V and 1.95 respectively.

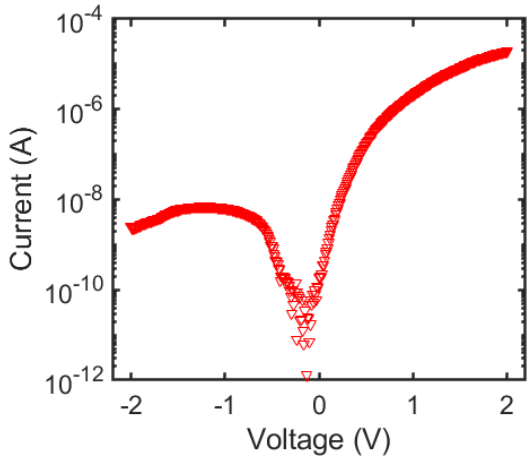


Fig. 5 ln I-V characteristics of heterojunction diode under consideration, showing rectifying behavior with turn-on voltage = 0.83 V for the determination of I_0, Φ_B and η.

2) Capacitance – Voltage (C-V) characteristics:

The measured C-V characteristic of the n-MoS_2/p-Si heterojunction diode is shown in Fig. 6(a). From Fig. 6(a), it is clear that n-MoS_2/p-Si heterojunction can be represented by the one-sided p^+-n abrupt junction. Thus the C–V characteristics of the heterojunction device are described by the following relation [19]:

$$\frac{A^2}{C^2} = \frac{2\left(V_{bi}+V-\frac{kT}{q}\right)}{qN_D\varepsilon_s}, \quad (2)$$

where N_D is the donor concentration of n-MoS_2, ε_s is the permittivity, V is the applied reverse bias voltage, and V_{bi} is the built-in potential. The plot of A^2/C^2 with applied bias V is shown in Fig. 6(b). Eq. (2) shows that the A^2/C^2 versus V should be a linear whose intercept with the applied voltage bias gives the value of − (V_{bi} – kT/q) as -1.58 [From Fig. 6(b)]. The slope (denoted as "m") is used to estimate the doping concentration N_D of MoS_2 by using the given relation:

$$N_D = \frac{2}{mq\varepsilon_s} \quad (3)$$

The value of V_{bi} calculated as 1.6 eV and the value of N_D has been calculated as 8.326 × 10^{14} cm^{-3}. The $\Phi_{B,\,CV}$ can be determined from the following relation:

$$\Phi_{B,CV} = V_{do} + V_n, \quad (4)$$

where V_{do} is the diffusion voltage ($V_{bi} + kT/q$) and $V_n = \left(\frac{kT}{q}\right)\ln(\frac{N_C}{N_D})$ is the depth of the Fermi level. The barrier height of n-MoS_2/p-Si heterojunction diodes from C-V characteristics has been estimated as 1.87 eV.

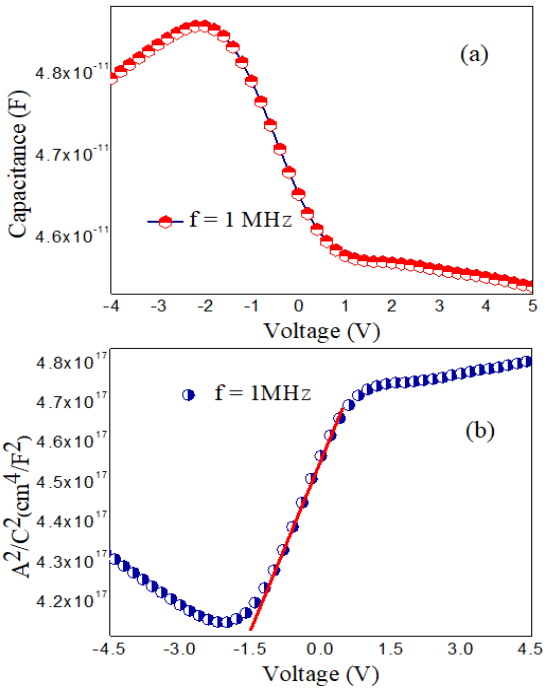


Fig. 6 (a) The C–V characteristics of n-MoS_2/p-Si heterojunction device; **(b)** The A^2/C^2 vs. V plot for the determination V_{bi} and N_D.

IV. CONCLUSION

The solution-based growth of MoS_2 QDs on p-Si substrates has been successfully demonstrated by a cost-effective facile colloidal approach for n-MoS_2/p-Si heterojunction devices applications. Excitation dependent PL spectra of MoS_2 show polydispersity of QDs dispersions and this behavior is useful in multi-colour imaging applications. The barrier height and doping concentration of n-MoS_2/p-Si heterojunction diode are found to be 1.87 eV and 8.326×10^{14} cm^{-3} calculated from the C–V measurement. The value of ϕ_B from I-V measurements is estimated to be 0.23 V. A large difference between barrier height value observed from C-V and I-V measurements is the due presence of large interface defects at the n-MoS_2/p-Si junction. The results are found to be promising for various electronics and optoelectronics device applications.

ACKNOWLEDGMENT

D. Somvanshi thanks the DST INSPIRE faculty award (DST/INSPIRE/04/2017/000147) of the Government of India for the financial support. The authors also thank Dr. Chandan Kumar and Prof. S. Jit, Department of Electronics Engineering, IITBHU (Varanasi) for the measurement of electrical characteristics of devices.

REFERENCES

[1] J. Y. Lee, J.-H. Shin, G.-H. Lee, and C.-H. Lee, "Two-Dimensional Semiconductor Optoelectronics Based on van der Waals Heterostructures," *Nanomaterials (Basel, Switzerland),* vol. 6, p. 193, 2016.

[2] H. Wang, C. Li, P. Fang, Z. Zhang, and J. Z. Zhang, "Synthesis, properties, and optoelectronic applications of two-dimensional MoS2 and MoS2-based heterostructures," *Chemical Society Reviews,* vol. 47, pp. 6101-6127, 2018.

[3] Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, and M. S. Strano, "Electronics and optoelectronics of two-dimensional transition metal dichalcogenides," *Nature Nanotechnology,* vol. 7, p. 699, 11/06/online 2012.

[4] M. Ye, D. Zhang, and Y. K. Yap, "Recent Advances in Electronic and Optoelectronic Devices Based on Two-Dimensional Transition Metal Dichalcogenides," *Electronics,* vol. 6, p. 43, 2017.

[5] D. Somvanshi and S. Jit, "Transition metal dichalcogenides based two-dimensional heterostructures for optoelectronic applications," in *2D Nanoscale Heterostructured Materials*, ed: Elsevier, 2020, pp. 125-149.

[6] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, "Atomically Thin MoS_2: A New Direct-Gap Semiconductor," *Physical Review Letters,* vol. 105, p. 136805, 09/24/ 2010.

[7] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, "Single-layer MoS2 transistors," *Nature Nanotechnology,* vol. 6, p. 147, 01/30/online 2011.

[8] W. Choi, M. Y. Cho, A. Konar, J. H. Lee, G.-B. Cha, S. C. Hong, *et al.*, "High-Detectivity Multilayer MoS2 Phototransistors with Spectral Response from Ultraviolet to Infrared," *Advanced Materials,* vol. 24, pp. 5832-5836, 2012.

[9] S. Ahmad and S. Mukherjee, "A Comparative Study of Electronic Properties of Bulk MoS2 and Its Monolayer Using DFT Technique: Application of Mechanical Strain on MoS2 Monolayer," *Graphene,* vol. 03, pp. 52-59, 01/01 2014.

[10] S. Mukherjee, R. Maiti, A. K. Katiyar, S. Das, and S. K. Ray, "Novel Colloidal MoS2 Quantum Dot Heterojunctions on Silicon Platforms for Multifunctional Optoelectronic Devices," *Scientific Reports,* vol. 6, p. 29016, 2016/06/30 2016.

[11] N. S. Arul and V. D. Nithya, "Molybdenum disulfide quantum dots: synthesis and applications," *RSC Advances,* vol. 6, pp. 65670-65682, 2016.

[12] V. Dhyani, P. Dwivedi, S. Dhanekar, and S. Das, "High-performance broadband photodetector based on MoS2/porous silicon heterojunction," *Applied Physics Letters,* vol. 111, p. 191107, 2017.

[13] H. Lin, C. Wang, J. Wu, Z. Xu, Y. Huang, and C. Zhang, "Colloidal synthesis of MoS2 quantum dots: size-dependent tunable photoluminescence and bioimaging," *New Journal of Chemistry,* vol. 39, pp. 8492-8497, 2015.

[14] Z. X. Gan, L. Z. Liu, H. Y. Wu, Y. L. Hao, Y. Shan, X. L. Wu, *et al.*, "Quantum confinement effects across two-dimensional planes in MoS2 quantum dots," *Applied Physics Letters,* vol. 106, p. 233113, 2015.

[15] H. Nan, Z. Wang, W. Wang, Z. Liang, Y. Lu, Q. Chen, *et al.*, "Strong Photoluminescence Enhancement of MoS2 through Defect Engineering and Oxygen Bonding," *ACS Nano,* vol. 8, pp. 5738-5745, 2014/06/24 2014.

[16] Z. Wang, L. Cao, Y. Ding, R. Shi, X. Wang, H. Lu, *et al.*, "One-step and green synthesis of nitrogen-doped carbon quantum dots for multifunctional electronics," *RSC Advances,* vol. 7, pp. 21969-21973, 2017.

[17] W. Yin, X. Bai, X. Zhang, J. Zhang, X. Gao, and W. W. Yu, "Multicolor Light-Emitting Diodes with MoS2 Quantum Dots," *Particle & Particle Systems Characterization,* vol. 36, p. 1800362, 2019.

[18] B. Li, L. Jiang, X. Li, P. Ran, P. Zuo, A. Wang, *et al.*, "Preparation of Monolayer MoS2 Quantum Dots using Temporally Shaped Femtosecond Laser Ablation of Bulk MoS2 Targets in Water," *Scientific Reports,* vol. 7, p. 11182, 2017/09/11 2017.

[19] D. Somvanshi and S. Jit, "Analysis of Temperature-Dependent Electrical Characteristics of n-ZnO Nanowires (NWs)/p-Si Heterojunction Diodes," *IEEE Transactions on Nanotechnology,* vol. 13, pp. 62-69, 2014.

Spherical Phase Method using Phase Center Determination of a Dual-Band Patch Antenna

Rintu Kumar Gayen, Payel Midya, Sutanu Ghosh, Kaushik Patra, Malay Gangopadhyay
Department of Electronics & Communication Engineeinrg
Institute of Engineering and Management,
Salt Lake, Kolkata - 700091, India
E-mail id: rintukrgayen@gmail.com, payel.midya@iemcal.com, sutanu.ghosh@iemcal.com,
kaushik.patra@iemcal.com, malay.ganguly@iemcal.com

***Abstract*— In this report, the spherical phase method using phase center determination of a dual-frequency rectangle loaded monopole patch antenna is presented. This rectangular tuning patch and dual individual wide bands can be achieved by further adjusting the dimensions and positions of this structure and phase center determination of a dual-frequency patch antenna and the distance between two-phase centers is calculated.**

Keywords*— *Phase Center, Dual Frequency Antenna, Monopole Patch Antenna, Rectangular Tuning Patch, Spherical Phase Method.

I. INTRODUCTION

The electrical position of the mean phase center offset (PCO) of an antenna plays an important role in various modern antenna engineering efforts. The issue is important in utilizations such as telemetry, imaging (radio telescopes, imaging radars), antenna arrays, ultra-wideband (UWB) applications with signal correlation and group delay distortion which are primarily attributed to phase center variations (PCV).

The phase center of a practical antenna differs in various azimuth and elevation angles. This is called phase center dispersion. The antenna phase center can only be kept comparatively stable for a definite range of directions. The concepts of average phase center, phase center offset and phase center variation are utilized to explain fully the characteristics of antenna phase response. The average phase center of an antenna is described as the mean location of the antenna phase center over a certain range of directions. The antenna reference point (ARP) is also defined as the geometric center of the antenna. The distinction between the average phase center and the antenna reference point is considered as the phase center offset (PCO). The offset between the phase center in a particular direction and the average phase center is described as phase center variation (PCV) in that direction.

Two methods of calculations are available in the literature to calculate the phase center of an antenna. One is based on the measurements in an anechoic chamber [1] and the other is mathematical calculation [2-10].

Several researchers have concentrated on the phase center determination techniques. In [3], the authors focused on the location of the phase centers for microwave antennas.The difficulty of traditional aperture techniques for the accurate description of phase centers using the second derivative method is discussed. In [4], two-point method is used to determining phase centers and its application to horn antennas is presented. An extension of the edge diffraction method to determine amplitude patterns is presented in [5] to predict the radiation phase center of pyramidal horn antennas. In [6], a novel differential phase method is utilized to measure the phase center location and the phase pattern for a 110 GHz corrugated horn. Phase center measurements based on the three-antenna method was presented in [7]. In [8], the calculation of the impact of the virtual phase centre effect on position system accuracy is presented. Time of arrival (TOA) method is presented in [9], Spherical phase method in global navigation satellite system (GNSS) requiring the exact position of antenna phase center is presented in [10] and comparison of different methods for determining the antenna phase center experimentally using a planar acquisition system in [11].

In the literature, many authors reported the phase center calculation of a single frequency antenna. But many applications need the use of dual-frequency patch antennas for which not much information is available for phase center calculation in the pieces of literature. In this letter, phase center determination of a dual-frequency rectangle loaded monopole patch antenna [12 - 13] with a compact radiator (10×23 mm^2) is presented using spherical phase method and the distance between two-phase centers is calculated.

II. ANTENNA DESIGN

Layout of the proposed dual-band monopole antenna is shown in Figure 1. Substrate for the antenna is chosen as FR4 with a substrate thickness of 1.6 mm with a relative permittivity of 4.4 and a loss tangent of 0.02. A coplanar waveguide (CPW) based feeding mechanism is utilized.

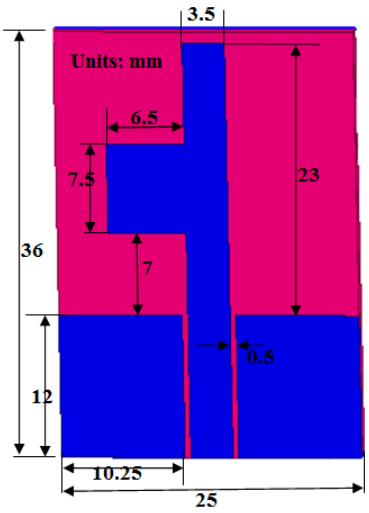


Figure 1: Configuration of dual frequency rectangle-loaded monopole patch antenna

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

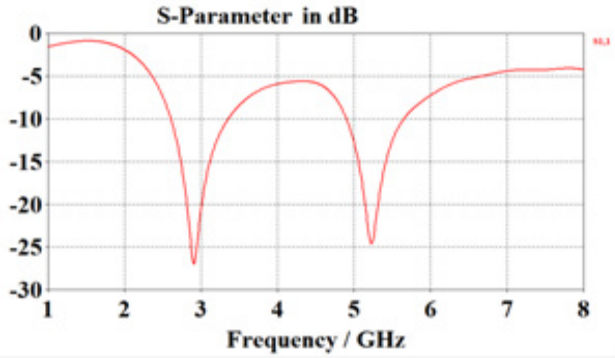


Figure 2. The magnitude in dB of reflection coefficient for proposed antenna.

A CPW line of characteristic impedance 50 Ω having a signal strip width of 3.5 mm and a gap of 0.5 mm between the signal strip and the coplanar ground is utilized. First, a straight monopole antenna is designed and simulated using CST Microwave Studio. Subsequently a rectangular tuning patch is added to the straight strip of antenna based on the modal technique [13]. This leads to the design of Fig. 1. By tuning the dimensions of the patch and its position along the straight strip, two resonant frequencies at about 2.9 and 5.2 GHz, respectively, are obtained, as shown in Figure 2.

III. Description of the Method

In this work a simple method of calculating the phase center based on practical antenna is presented. The spherical phase method [10] is utilized starting with the spherical phase expansion from its origin located at point PCO (x, y, z). A fixed constant "c" is added. Azimuth and elevation angles are represented as θ and φ, while wavelength is λ. Azimuth and elevation angles values are obtained from the simulation by CST software.

$$Phase(\phi,\theta)=\frac{2\pi}{\lambda}\begin{pmatrix}x\cdot\cos\phi\sin\theta+\\ y\cdot\sin\phi\sin\theta+\\ z\cdot\cos\theta\end{pmatrix}+c$$

or in matrix

$$\text{notation } F(\phi,\theta)=\frac{2\pi}{\lambda}\left[\cos\phi\sin\theta\quad\sin\phi\sin\theta\quad\cos\theta\right]\cdot\begin{bmatrix}x\\y\\z\\c\end{bmatrix}$$

(1a)

$$F=M\cdot P \qquad (1b)$$

To find phase center P we use eq. (1b), hence:

$$P=M^{-1}\cdot F$$

$$\begin{bmatrix}x\\y\\z\\c\end{bmatrix}=\frac{\lambda}{2\pi}\left[\cos\phi\sin\theta\quad\sin\phi\sin\theta\quad\cos\theta\right]^{-1}\cdot F(\phi,\theta)$$

Assembling a solution matrix with measurements obtained from samples of the phase pattern of the antenna, we obtain a mean value of PCO as.

$$\begin{bmatrix}x\\y\\z\\c\end{bmatrix}=\frac{1}{2\pi}\bullet\begin{bmatrix}\cos\phi_1\sin\theta_1 & \sin\phi_1\sin\theta_1 & \cos\theta_1 & 1\\ \cos\phi_2\sin\theta_1 & \sin\phi_2\sin\theta_1 & \cos\theta_1 & 1\\ \vdots & \vdots & \vdots & \vdots\\ \cos\phi_N\sin\theta_1 & \sin\phi_N\sin\theta_1 & \cos\theta_1 & 1\\ \cos\phi_1\sin\theta_2 & \sin\phi_1\sin\theta_2 & \cos\theta_2 & 1\\ \vdots & \vdots & \vdots & \vdots\\ \cos\phi_N\sin\theta_K & \sin\phi_N\sin\theta_K & \cos\theta_K & 1\end{bmatrix}\bullet$$

$$\begin{bmatrix}\varphi_1(\phi_1,\theta_1)\cdot\lambda\\ \varphi_1(\phi_2,\theta_1)\cdot\lambda\\ \vdots\\ \varphi_N(\phi_N,\theta_1)\cdot\lambda\\ \varphi_{N+1}(\phi_1,\theta_2)\cdot\lambda\\ \vdots\\ \varphi_{N\cdot K}(\phi_N,\theta_K)\cdot\lambda\end{bmatrix}$$

$$where:\quad \phi_i=0:\frac{2\pi}{N}:2\pi,\quad \theta_i=0:\frac{\pi}{2K}:\frac{\pi}{2}$$

(upper hemisphere)

To obtain the mean PCO solution for all frequencies at one go using (3),more columns can be added to the right representing mathematical calculations made at different frequencies. After the mean PCO is determined, we can obtain an ideal spherical phase pattern anchored at mean PCO by multiplying back with the spherical phase expansion matrix. The difference residual between the ideal and calculated phase pattern will be the phase center variation (PCV).

IV. Results and Discussion

With the dimensions as shown in figure 1, the simulated magnitude in dB of the reflection coefficient of the proposed antenna obtained from CST software is shown in figure 2. The proposed antenna, on simulation, gives two resonant modes at about 2.9 and 5.2 GHz, respectively, as shown in Figure 2.

The CST simulated far-field radiation pattern at 2.9 GHz respectively, is plotted in figure 3. From figure 3, we find the phase center value (-0.391147, -1.33637, 0.0524673) at 2.9 GHz.

Similarly, the simulated phase center at 5.2233 GHz respectively, are plotted in Fig. 4.From Fig. 4 we find the phase center value (0.188681, -0.590922, -2.06147) at 5.2 GHz.

The distance between two phase centers is *d*

$$d=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}$$

$$d=2.3132228599943\text{ mm} \qquad (1)$$

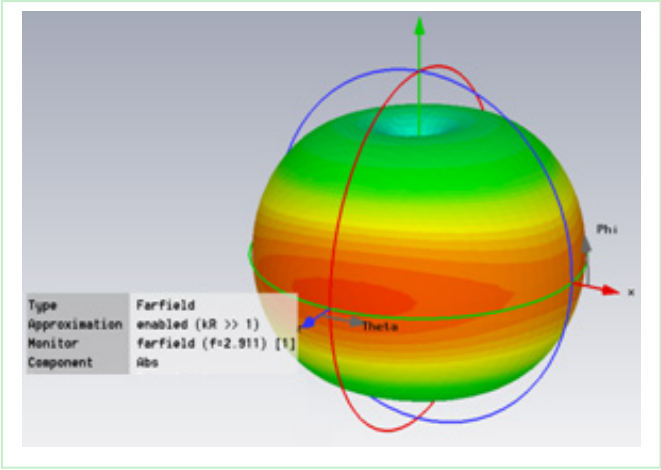


Fig. 3. 3D radiation pattern of dual frequency patch antenna (at f = 2.9 GHz).

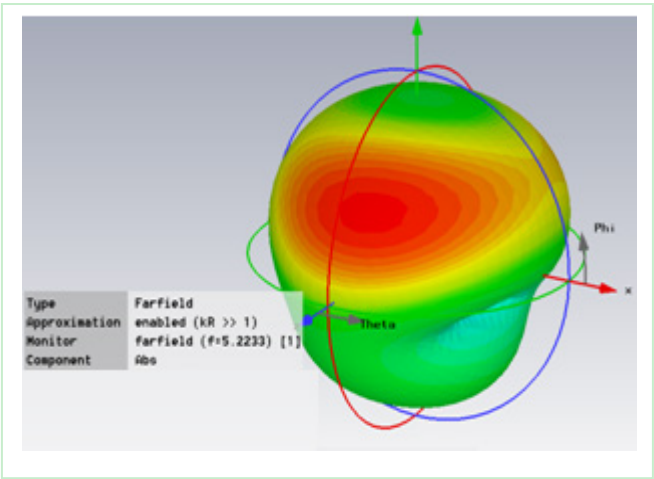


Fig. 4. 3D radiation pattern of dual frequency patch antenna using (at f = 5.2 GHz).

V. Conclusion

This paper intends to present a simple method of computing phase center based on a real antenna, using the spherical phase method. For the taken dimensions of the antenna by CST simulation, two resonant frequencies are obtained at 2.9 GHz and 5.2 GHz. The far-field radiation patterns at 2.9 GHz and 5.2 GHz are shown in figure 3 & 4 respectively. Obtained co-ordinates of their phase centers from far-field are (-0.391147, -1.33637, 0.0524673) and (0.188681, -0.590922, -2.06147), respectively. All these data are obtained on simulation of the proposed antenna on CST software. The distance between the two-phase centres is determined with the help of mathematical formula which comes out to be 2.3132228599943 mm.

References

[1] Hu. Y. Z., Li. Z., Gang. O. and Bin. Z., "Research On Antenna Phase Center Anechoic Chamber Calibration Method", International Conference on Microwave and Millimeter Wave Technology (ICMMT), 2010.

[2] Kunysz. W., "Antenna Phase Center Effects and Measurements in GNSS Ranging Applications", International Symposium on Antenna Technology and Applied Electromagnetics & the American Electromagnetics Conference (ANTEM-AMEREM), 2010.

[3] D.Carter, "Phase Centers of microwave antennas", IRE Trans. Antennas and Prop., Vol AP-4, pp 597-600, Oct. 1956.

[4] Y.Y.Hu, "A method of determining phase centers and its applications to electromagnetic horns", J. Franklin Inst., Vol.271, pp 31-39, Jan. 1961.

[5] M. Teichman, "Determination of horn antenna phase centre by edgediffraction theory", IEEE Trans. Aerospace Eelctr. Systems, Vol. AES-9, pp 875-882, Nov. 1973.

[6] J. Thovinen, A. Lehto, A. Raisanenn, "Differential Phase Method", Int.Symp. of Antennas and Propag. Society, Vol. 3, pp 1298-1301, Ma 7-11, 1990.

[7] T.W.Hertel,"Phase Center based on the three antenna method", IEEEInt. Symp. of Antennas and Propag. Scoiety, Vol. 3, pp 816-819, June22-27, 2003.

[8] Y. Yashcheshyn, M.Bury, K.Kurek and P.Bajurko, "Evaluation of the impact of the virtual phase center effect on the accuracy of the positioning system", 3rd European Conf. on Antennas and Prop., pp. 2930- 933, March 23-27, 2009.

[9] R.C. Jones, "New calculus for the treatment of optical systems," J.Opt.Soc. Am. 31,pp. 448-493, 1941.

[10] Kumar. A., Ansari E. and Yedukondalu. K., "Improved Phase Center Estimation for GNSS Patch Antenna", IEEE Transactions on Antennas and Propagation, Vol. 61, no. 4, April 2013.

[11] P. Padilla1, J. M. Fernandez, J. L. Padilla, Gonzalo E-D, ManuelS-C, and B. Galocha, "Comparison of different methods for the experimental antenna phase center determination using a planar acquisition system", Progress In Electromagnetics Research, Vol. 135, pp. 331-346, 2013.

[12] Chung. Y., Jeon. S. and Itoh. T., "High Isolation Dual-Polarized Patch Antenna using Integrated Defected Ground Structure", IEEE Microwave and Wireless Components Letters, Vol. 14, no. 1, January 2004.

[13] H. Chen, X. Yang, Y.-Z. Yin, J.-J. Wu and Y.-M. Cai, "Tri-band rectangle-loaded monopole antenna with inverted-L slot for WLAN/WiMAX applications", Electronics Letters, Vol. 49 No. 20 pp. 1261–1262, 2013.

[14] Wong, K.-L., and Huang, C.-H.: 'Printed loop antenna with a perpendicularfeed for penta-band mobile phone application', IEEE Trans. Antennas Propag., 56, (7), pp. 2138–2141, 2008.

[15] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*

[16] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[17] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[18] K. Elissa, "Title of paper if known," unpublished.

[19] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[20] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[21] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

Behaviour Shockley and Sakurai Models in 7nm FinFet

Mohammad Anees
Memory Design Group
Xilinx Pvt. Ltd
Hyderabad, India
manees@xilinx.com

Kumar Rahul
Memory Design Group
Xilinx Pvt. Ltd.
Hyderabad, India
kumarr@xilinx.com

Sourabh Aditya Swarnkar
Memory Design Group
Xilinx Pvt. Ltd.
Hyderabad, India
sourabh@xilinx.com

Santosh Yachareni
Memory Design Group
Xilinx Pvt. Ltd
Hyderabad, India
santoshy@xilinx.com

Abstract— As we move towards lower nodes, device dimensions are changing and becoming more complex. With FinFet's elevated channel, gate voltage is applied from all 3 directions. The gate electric field is no longer unidirectional. These add complexity in the behavior of current through the channel. Traditional Shockley model which had square law dependency on the over drive voltage in saturation region will no longer hold good. To overcome this shortcoming of the model alpha power model was introduced by Sakurai-Newton [1]. This accurately modeled the transistor behavior in submicron device in saturation region. According to this model the device will not have square dependency in saturation region. But will have power reduced from two to one as node shrinks. As we were scaling the nodes, there wasn't any significant scaling in the voltage. As the velocity saturation becomes more dominant alpha reaches one. Models designed for the planar devices and are not in sync with the FinFet. Proposed paper describes the behavior of the MOS in 7nm across different voltage condition and comparison of the current behavior predicted from the Shockley's square law model and Sakurai-Newton's alpha power model in FinFets.

Keywords— Finfets, MOSFET (Metal Oxide Semiconductor Field Effect Transistors), Ids (Drain to Source current), Vds (Drain to Source Voltage), Vgs (Gate to Source Voltage), Shockley & Sakurai Models, Velocity Saturation, Alpha-Power Law Model, GDTC (Gate Direct Tunneling Current), GIDL (Gate-Induced Drain Leakage), DIBL (Drain-Induced Barrier Lowering), NBTI (Negative Bias Temperature Instability)

I. INTRODUCTION

As the technology scales to more advanced nodes there are many second order effects which was ignored in long channel length becomes more prominent. Constant voltage scaling reached its limit because of high degree of velocity saturation and other effects such hot carrier related problems. With adoption of constant field scaling with technology, the voltage was decreased accordingly. Now this too reaching its limit. Voltage scaling is not happening as aggressively as it was earlier. As MOS dimensions are scaled, device performance deviates from the simple models of large devices traditionally used by circuit designers [2]. In making rough calculations during initial stages of design, most second-order effects can be neglected in properly scaled technology. As the feature size of MOSFET continue to decrease (<100nm), many new physical effects are being observed due to the limitation of technologies [3]. Different effects like Gate Direct Tunneling Current (GDTC), Gate-Induced Drain Leakage (GIDL), Drain-Induced Barrier Lowering (DIBL), Negative Bias Temperature Instability (NBTI) etc. affects the operation of MOSFETs [4]. However, the effects of the increased normal (gate) and tangential (drain) electric field cannot be neglected without considerable error when dealing with modern device sizes [5]. With high electric field at smaller dimensions, increases velocity saturation effect in the device. In FinFet era the device has become more complex with elevated channel makes the device behavior complex [6]. This paper discusses about the conventional Shockley modeling and its shortcoming. Later the alpha power model described by Sakurai and application of the model in the FinFets.

II. MOS MODEL

A. Shockley & Sakurai-Newton models

Following shows the Shockley's model for the current in the MOSFET [1].

$$I_D = \begin{cases} 0 & (V_{GS} \leqslant V_{TH}\text{: cutoff region}) \\ K\{(V_{GS}-V_{TH})V_{DS} - 0.5V_{DS}^2\} & (V_{DS} < V_{DSAT}\text{: linear region}) \\ 0.5K(V_{GS}-V_{TH})^2 & (V_{DS} \geqslant V_{DSAT}\text{: saturation region}) \end{cases}$$

Fig.1

Fig. 1 shows the square law [8] model for the MOS and as per the equations in saturation region the current is increasing in square of over drive voltage. But at sub-micron channel length because of high electric field at the gate velocity saturation becomes more dominant. As a result the current doesn't increase quadratically with overdrive voltage in saturation region. This was captured by Sakurai in alpha power model. The Alpha-Power Law MOSFET Model is the most widely utilized compact drain current model due to its simple mathematical form and high degree of accuracy[7]. As per the model current won't have any square law dependencies. Because of velocity saturation caused due to the high electric field in the channel. The increase in current can be varied between 1 and 2 based on the severity of the velocity saturation. This means, bigger the channel length greater the alpha value. Alpha value cannot go back to the square as we move towards the lower nodes. And it is described as follows [1].

$$Id = \begin{cases} 0 & (Vgs \leq Vth\text{: Cutoff region}) \\ \left(\frac{Ido'}{Vdo'}\right) * Vds & (Vds < Vdo'\text{: Triode region}) \\ Ido' & (Vds \geq Vdo'\text{: Saturation region}) \end{cases}$$

Where,

$$Ido' = Id0 * \left(\frac{Vgs-Vth}{Vdd-Vth}\right)^{\propto} = \frac{W}{Leff} * Pc * (Vgs-Vth)^{\propto}$$

$$Vdo' = Vd0 * \left(\frac{Vgs-Vth}{Vdd-Vth}\right)^{\propto/2} = Pv * (Vgs-Vth)^{\propto/2}$$

Fig. 2

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Alpha-Power Law proposed by Sakurai and Newton [1] includes the carrier velocity saturation effect comparable with the SPICE MOS level-3 model [9].

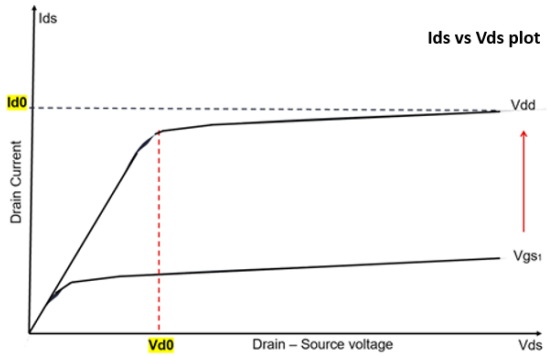


Fig. 3

The Sakurai model [1] described in the Fig. 2 & Fig. 3 have following parameters. Vth (Threshold voltage), α (velocity saturation index), Vdo (Drain saturation voltage at Vgs =Vdd), Ido (Drain current at Vgs =Vdd). This model was targeted for the behavior in saturation region and hence this is not modeled for linear region or subthreshold region analysis.

Analysis includes how well we can predict the current behavior of FinFet using Shockley and Sakurai-Newton model. And the behavior of the current in FinFets, as it is having complex structure compared to planar MOSFETs.

III. FINFET BEHAVIOR

FinFets have three dimensional structure which has a thin silicon body perpendicular to the plane of the wafer. The channel of the FinFet is wrapped by the gate in three directions by the three dimensional structure of the FinFet device [10]. And the channel is completely depleted. Current will flow in three ways since the channel formed is due to elevated gate. Following fig. 4 [11] shows physical comparison between the FinFet and planar MOSFET. With FinFets the threshold voltage can be changed by more tapering the elevated channel and also can be controlled by costly work function modification which we use in more conventional devices. The two main working modes for FinFets are Shorted-Gate (SG) mode and Independent-Gate (IG) mode [12]. For our analysis we are using shorted gate FinFets.

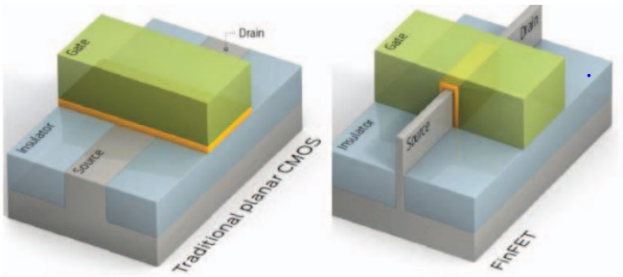


Fig. 4

FinFets have added more variables into the behavior of the transistors in comparison to the planar MOSFET [13]. And saturation phenomenon in FinFets are not like that of the planar devices [14]. In the following section we will first observe the behavior of FinFet in different Gate and Drain voltage scenarios. We will check the response of the FinFet by probing the current through it.

A. Modeling current and voltage behavior in FinFet

Modeling current in MOSFET is very useful for designer to do a quick calculation. With actual models used in spice are having large number of variables and more complicated equation to predict the behavior of the MOSFET. But using such equations by designer to design any circuits is not feasible. To model a device, we need to check its behavior under different circumstances. The reduced mobility at high electric field cause significant reduction in Idsat (Saturation current) and transconductance gm [15]. In following section, we will see how well the traditional Shockley's square law model and Sakurai's alpha power model predicts the behavior of FinFets.

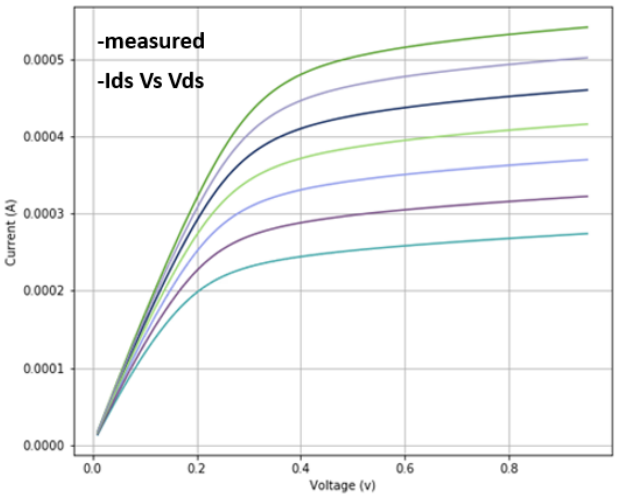


Fig. 5

Fig. 5 is the Ids Vs Vds characteristics of 7nm FinFet device. Vgs is chosen between 650mv to 950mv and Vds is varied from 0 to 950mv. As it is evident from above figure 5 is that we don't see any square increase in current with respect to over drive voltage.

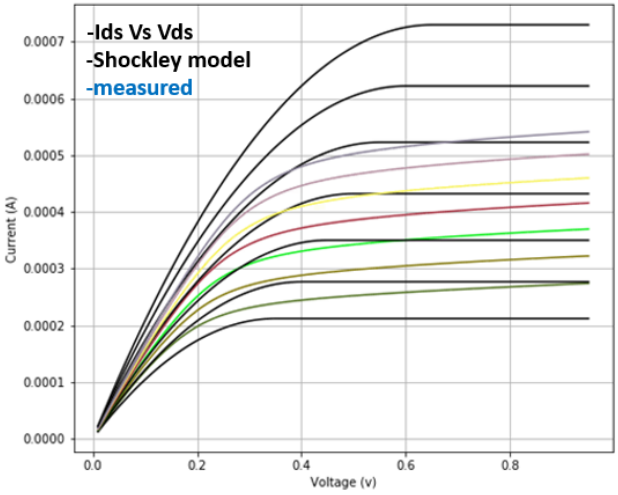


Fig. 6

In Fig. 6 we have combined the Shockley's square law model result with the FinFet Ids vs Vds characteristics. Black color line is calculated using Square law model and the colored one is the FinFet behavior. It is very clear that the current predicted by the Square law model is nearly 40% higher as compared to the actual one. And we can also see

the square law model deviates as the over drive voltage increases. And at the high over drive voltage model predicts more current than the actual. This is because the square law model did not account for the effect of velocity saturation.

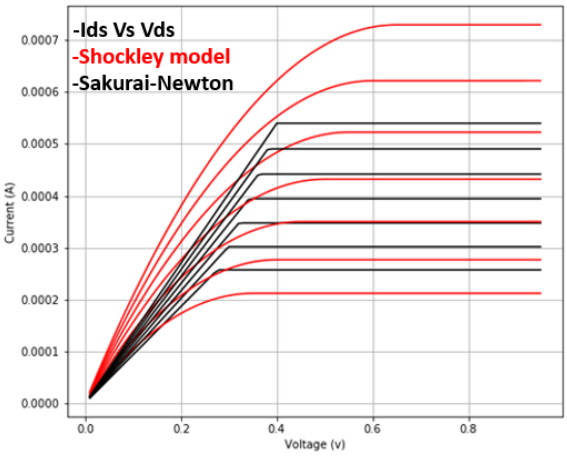


Fig. 7

Fig. 7 shows the Shockley's square law model in red color and Sakurai-Newton's alpha power model in black color. And we can see the difference in the current profile. As the Shockley square law model failed to consider the impact of the velocity saturation which produces early saturation of current. And the increment in current per over drive voltage in saturation region is not increasing quadratically. Sakurai have considered the velocity saturation effect and the same has been modeled.

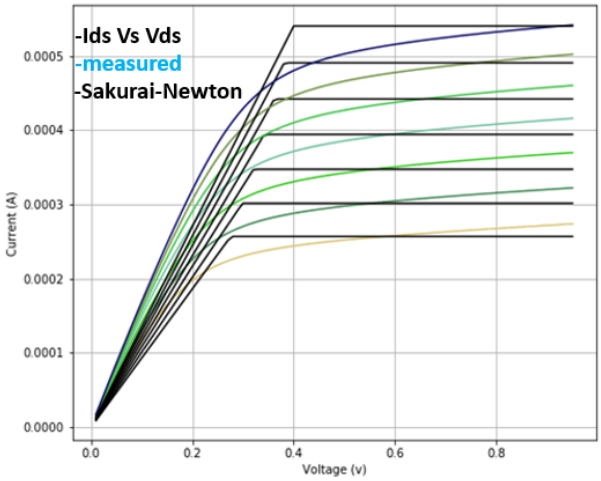


Fig. 8

In Fig. 8 black color line shows the alpha power model-based prediction and the colored is the measured. Current is modeled more closely to measured value compared to the Shockley's square law model. But as the Vgs increases, difference in the prediction and measured value deviates more.

To understand the behavior of the FinFet in saturation region. Let's consider the Ids Vs Vgs plot. Vds is maintained at 950mv and Vgs is varied from 0 to 950mv. Threshold voltage of the transistor under test is ~300mv. In this setup the FinFet is always in saturation region.

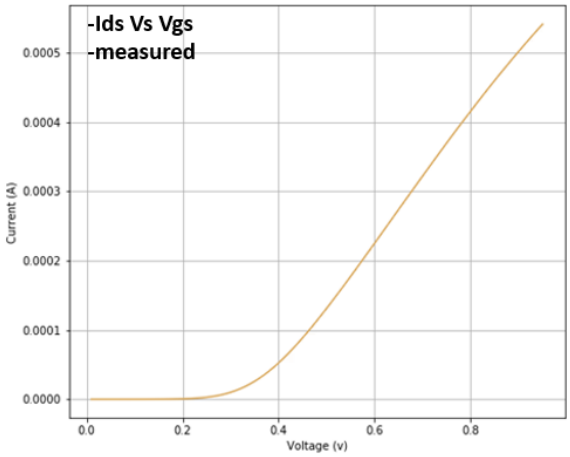


Fig. 9

Fig. 9 Is Ids Vs Vgs plot. But upon taking Log_{10} of both the axis gives the rate at which the current is changing per increase in voltage. If the current is increasing at same rate, then the curve should be ideally same throughout the voltage.

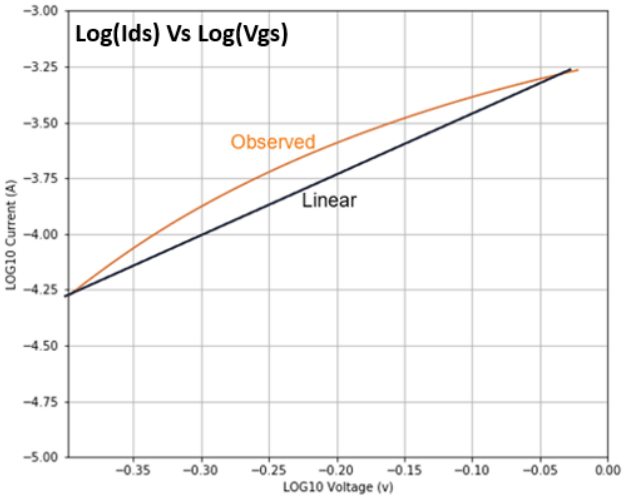


Fig. 10

Fig.10 shows the Fig. 9 plot with Log_{10} taken on both the axis. Here the gate to source voltage varies from 400mv to 950mv. And same is plotted in Log scale.

As seen in above Fig. 10 that the line is not linear and tend to taper at higher voltages. This means that rate at which current increases for constant increase in voltage is not same. Conventional way of calculating the alpha is to take the slope of the Fig. 10 that would give us the rate at which current increases.

$$y = f(x) \quad – \quad (1)$$

$$f(x) = x^2 \quad – \quad (2)$$

$$y_1 = x_1^2 \quad – \quad (3)$$

$$y_2 = x_2^2 \quad – \quad (4)$$

Taking log on both side of 3 and 4. Taking slope at y_1, y_2 and x_1, x_2 will yield 2. And this is the rate at which the y is changing with respect to x. Applying same to Fig. 10 as shown in Fig. 11 [1].

$$\alpha = \frac{Log(\frac{Id1}{Id2})}{Log((Vg1 - Vth)/(Vg2 - Vth))}$$

Fig. 11

Fig. 11 yields different alpha value for different Id1, Id2, Vg1, Vg2 values. In Fig. 8 the alpha value in the model is iteratively changed to have a close match with the measured values. And the value of alpha for which the model is close to measured results is 1.25.

IV. CONCLUSIONS

The Sakurai model is well aligned with the FinFet as compared to Shockley's square law model. But because of alpha value not constant even when the device in saturation extracting the accurate alpha value with few sample points is difficult. Alpha power model can be enhanced to cover the channel length modulation in sub 45nm and FinFet devices. At higher gate voltages the rate of change of current deceases which can be modeled in the alpha power model. With technology scaling becoming aggressive and with new device structures it becomes important to model the MOS behavior in simple terms for designer. This will help designers to get the initial designs more accurate with simplified models.

REFERENCES

[1] T. Sakurai and A. R. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," in IEEE Journal of Solid-State Circuits, vol. 25, no. 2, pp. 584-594, April 1990, doi: 10.1109/4.52187.

[2] H. Shichman and D. A. Hodges, "Modeling and simulation of insulated-gate field-effect transistor switching circuits," in IEEE Journal of Solid-State Circuits, vol. 3, no. 3, pp. 285-289, Sept. 1968, doi: 10.1109/JSSC.1968.1049902.

[3] J. Samanta and B. P. De, "Comprehensive analysis of delay in UDSM CMOS circuits," 2011 International Conference on Electronics, Communication and Computing Technologies, Pauls Nagar, 2011, pp. 29-32, doi: 10.1109/ICECCT.2011.6077064.

[4] A. K. Singh, J. Samanta and J. Bhaumik, "Modified I–V model for delay analysis of UDSM CMOS circuits," 2012 International Conference on Communications, Devices and Intelligent Systems (CODIS), Kolkata, 2012, pp. 357-360, doi: 10.1109/CODIS.2012.6422212.

[5] Y. El-Mansy, "MOS Device and Technology Constraints in VLSI," in IEEE Journal of Solid-State Circuits, vol. 17, no. 2, pp. 197-203, April 1982, doi: 10.1109/JSSC.1982.1051716.

[6] Hyunsang Hwang, Jung-Suk Goo, Hoyup Kwon and Hyungsoon Shin, "Impact of velocity saturation region on nMOSFET's hot carrier reliability at elevated temperatures," Proceedings of 1995 IEEE International Reliability Physics Symposium, Las Vegas, NV, 1995, pp. 48-50, doi: 10.1109/RELPHY.1995.513652.

[7] K. A. Bowman, B. L. Austin, J. C. Eble, X. Tang and J. D. Meindl, "A physical alpha-power law MOSFET model," Proceedings. 1999 International Symposium on Low Power Electronics and Design (Cat. No.99TH8477), San Diego, CA, USA, 1999, pp. 218-222.

[8] K. Jeppson, "A learning tool MOSFET model: A stepping-stone from the square-law model to BSIM4," 2013 23rd International Workshop on Power and Timing Modeling, Optimization and Simulation (PATMOS), Karlsruhe, 2013, pp. 39-44, doi: 10.1109/PATMOS.2013.6662153.

[9] G. Angelov and M. Hristov, "SPICE modeling of MOSFETs in deep submicron," 27th International Spring Seminar on Electronics Technology: Meeting the Challenges of Electronics Technology Progress, 2004., Bankya, Bulgaria, 2004, pp. 257-262 vol.2, doi: 10.1109/ISSE.2004.1490430.

[10] L. Dileshwar Rao, S. Dixit, K. Pachkor and M. Aarthy, "FinFET based adiabatic logic design for low power applications," 2017 International conference on Microelectronic Devices, Circuits and Systems (ICMDCS), Vellore, 2017, pp. 1-6, doi: 10.1109/ICMDCS.2017.8211535.

[11] M. Chi, "FinFET technology: Overview and status at 14nm node and beyond," 2016 China Semiconductor Technology International Conference (CSTIC), Shanghai, 2016, pp. 1-3, doi: 10.1109/CSTIC.2016.7464027.

[12] D. Hisamoto et al., "FinFET-a self-aligned double-gate MOSFET scalable to 20 nm," in IEEE Transactions on Electron Devices, vol. 47, no. 12, pp. 2320-2325, Dec. 2000, doi: 10.1109/16.887014.

[13] L. Dileshwar Rao, S. Dixit, K. Pachkor and M. Aarthy, "FinFET based adiabatic logic design for low power applications," 2017 International conference on Microelectronic Devices, Circuits and Systems (ICMDCS), Vellore, 2017, pp. 1-6, doi: 10.1109/ICMDCS.2017.8211535.

[14] S. Banchhor, K. D. Kumar, A. Dwivedi and B. Anand, "A New Aspect of Saturation Phenomenon in FinFETs and Its Implication on Analog Circuits," in IEEE Transactions on Electron Devices, vol. 66, no. 7, pp. 2863-2868, July 2019, doi: 10.1109/TED.2019.2914867.

[15] S. L. Garverick and C. G. Sodini, "A simple model for scaled MOS transistors that includes field-dependent mobility," in IEEE Journal of Solid-State Circuits, vol. 22, no. 1, pp. 111-114, Feb. 1987, doi: 10.1109/JSSC.1987.1052681.

Ripple Correlation Control Maximum Power Point Tracking for Battery Operated PV Systems: A Comparative analysis

Pankaj Sahu[1]
Research Scholar (Ph.D.)
Dept. of ECE, SoET,
BML, Munjal University,
Gurugram, India
Email: pankaj.sahu@yahoo.co.in

Ashish Sharma[2]
Student,
Dept. of ECE, SoET,
BML, Munjal University,
Gurugram, India

Dr. Rajiv Dey[3]
Assistant Professor,
Dept. of ECE, SoET,
BML, Munjal University,
Gurugram, India

***Abstract*—This paper presents a comparative analysis of Ripple Correlation Control (RCC) with conventional Maximum Power Point Tracking (MPPT) methods such as Perturb & Observe (P&O) and Incremental Conductance (IC). Photovoltaic (PV) array along with the buck converter and its MPPT control have been simulated for all three MPPT methods at different solar radiation levels. From the simulation results, it can be observed that P&O is much affected with slow tracking and oscillations, and IC MPPT performs better than P&O in terms of tracking, but not for ripples. Simulation results verifies that RCC MPPT is capable of solving both these problems effectively and RCC performance is better than both P&O and IC MPPT methods.**

Keywords—PV, MPPT, RCC, HC, P&O, IC.

I. INTRODUCTION

Sustainable energy resources are always in high demand due of the fast and continuous depletion of fossil fuels and oil resources. Among the available green energy resources, solar energy has emerged as one of the most impactful renewable energy resource which has the potential to replace conventional energy sources. Solar energy is a free and constant source of renewable energy which is harvested using environment-friendly photovoltaic systems. However, it also suffers from some drawbacks e.g. during the day time, level of solar insolation is non uniform, moreover, increase in the environmental temperature degrades the power transfer to the load. Therefore, the power generated by solar panel is not same all the time and due to this each and every time the Maximum Power Point (MPP) of photovoltaic system gets changed [1]. Solar panel can deliver maximum power to the load for a certain value of voltage V_{MP} and current I_{MP}. This point is known as Maximum Power Point (MPP) as shown in Figure 2.

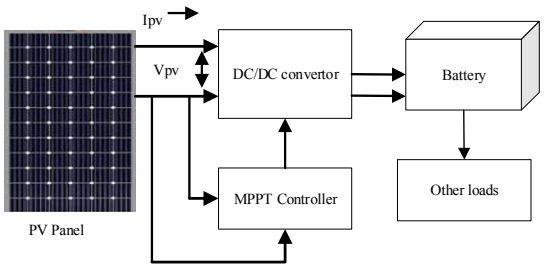


Fig. 1. Single stage battery operated PV system.

Since, the output power generated per watt by solar panel is too high, it is reasonable to improve its performance via power converters. These converters have many advantages in solar PV system like it can provide isolation between PV panel and load, extract maximum power. Moreover, converters are effective means for implementing MPPT technology [19]. MPPT algorithms are integrated with power electronic converter which can deliver maximum power by controlling its duty cycle. There are two types of available charge controllers; (i) Pulse Width Modulation (PWM) and (ii) Maximum Power Point Tracking (MPPT) [3]. In Photovoltaic systems MPPT based charge controller is preferred over PWM as a typical PWM charge controller can only be able to regulate the output voltage of a Photovoltaic (PV) array, not the current, but in MPPT controller both the voltage and current can be regulated. The block diagram of a battery operated PV system is shown in Figure 1.

Many research works have been done on MPPT in the last few decades such as conventional methods like Perturb and Observe [1-10], Incremental Conductance [11-16], Hill Climbing [17-19] have been deployed and a large number of modifications for these techniques has been proposed in the literature. The main shortcomings of P&O, HC & IC is that these methods can only track the maximum power when the irradiation conditions are uniform and it fails to track MPP when there is a partial shading or cloudy environmental condition. Further, these methods suffer from poor convergence, slow tracking speed, and high steady state oscillations (ripples). Therefore, to track the MPP under partial shading conditions conventional methods need to be modified. Due to which different modified versions of P&O have been proposed in the literature such as Adaptive P&O with fixed step size, variable step size, etc. [2-3]. Additionally, Soft Computing/Evolutionary algorithms [19] based approaches have also been emerged along with the conventional MPPT methods such as particle swarm optimization based MPPT algorithm [20] which is useful in finding the global MPP when there are multiple PV array with large area. These methods have numerous advantages like (i) ability to handle non linearity (ii) wide exploration in search space and (iii) coherent skill to reach global optimal regions these methods is considered to be a prime choice for non-linear optimization. Even-though the above alterations have enhanced the performance of MPPT charge controllers however, it is not sufficient enough for all the environmental conditions [6, 12]. This gives a motivation to the researchers to find some alternative ways to improve the performance of MPPT charge controllers under partial shading conditions.

On the other hand, new era of MPPT algorithms in PV system have emerged when evolutionary algorithm was first applied for MPP tracking [21-22]. Due to its ability of handle

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

nonlinear problems it has opened up new scopes of opportunities. Among these Ripple Correlation Control (RCC) [22] method has been well received by the research community because of its simplicity, robustness and ripple removal characteristics. Ripple Correlation Control (RCC) [21-26] provides a solution to the steady state oscillations, which is mainly caused by the internal ripples present in the system due to the presence of power electronic circuitry. Since, P&O, IC, and HC methods do not consider these ripples as internal perturbations, due to which the algorithm consider them as external perturbations. As a result the output voltage and current oscillates around MPP, in turn power also oscillates. These oscillations are harmful for the battery, as it reduces its operating life span.

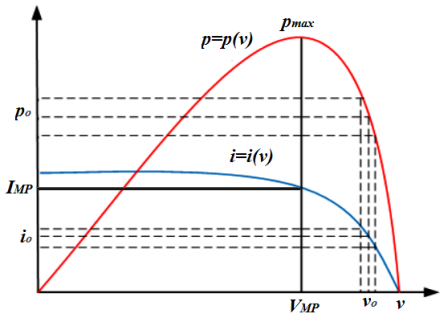


Fig. 2. Current and power of the PV panels versus voltage.

In this work, RCC MPPT with buck converter for battery connected load has been implemented and its simulation results have been compared with P&O and IC MPPT methods. The organization of the paper is as follows; Section I gives introduction of MPPT and various existing methods available in literature. Section II provides mathematical formulation of P&O, IC and RCC method, along with formulation of buck converter used in the simulation. In section III, parameters specifications are given along with photovoltaic system diagram. Section IV is simulation results and discussion, followed by conclusion in section V.

II. BACK GROUND OF THE CONVENTIONAL MPPT METHODS

In this section, a brief overview of the conventional MPPT methods along with detailed explanation of RCC control algorithm has been discussed.

A. Perturb & Observe (P & O)

The P&O Method [1-10] utilizes additional perturbations of current or voltage array to check if the system has achieved the nominal value of voltage or current. If the power output increases when voltage is changed in a particular direction of perturbation then it means that MPP will be obtained in that particular direction of perturbation and if the power output decreases for the same voltage perturbation then the MPP can be found by reversing the direction of perturbation. Although this method is easy to implement and is cost effective the overall system is never able to attain stability because apart from adding external perturbation to it the perturbation due to the environmental changes and the inherent generated perturbation affects its stability.

This method works on concept of continuous observations of PV array output power for the input perturbations, due to changing current or voltage of PV array. The algorithm continuously modifies the reference voltage or current based on the previous value of power until reaches the MPP [6, 7]. When $dP/dV > 0$ and the operating voltage of PV array is perturbed in a specific direction, it known that perturbation moves the operating point of PV array to the MPP. P&O method will then continue to perturb the PV voltage in the same direction. When $dP/dV < 0$, the perturbation moves the operating point of PV array away from the MPP and the P&O method reverses the direction of the perturbation [8, 9].

B. Incremental Conductance (IC)

Another algorithm to locate the MPP is Incremental Conductance (INC) [11-16]. In this algorithm a relationship between power and voltage is established where ideally the derivative of power with voltage is zero. This algorithm has both hardware and software complexity and is also expensive. The main drawback of this algorithm is that it increases the computation time of MPPT

$$\frac{dP}{dV} = \frac{d(VI)}{dV} = I\frac{dI}{dV} + V\frac{dI}{dV} = I + V\frac{dI}{dV} \quad . \tag{1}$$

Equation 1 represents the derivative of PV output power with voltage, the PV system will operate at MPP when this equation is equal to zero i.e.

$$\frac{dP}{dV} = 0 \Rightarrow I + V\frac{dI}{dV} = 0 \Rightarrow -\frac{I}{V} = \frac{dI}{dV} . \tag{2}$$

Comparing the instantaneous change in conductance, dI/dV and instantaneous conductance of PV array I/V in equation 2, the position of operating point in relation to maximum power point can be revealed. Equation 3 shows the position of operating point at different values of dP/dV

$$\begin{cases} \frac{dP}{dV} > 0, for\ V < V_{MPP} \\ \frac{dP}{dV} = 0, for\ V = V_{MPP} \\ \frac{dP}{dV} < 0, for\ V > V_{MPP} \end{cases} . \tag{3}$$

C. Ripple Correlation Control (RCC)

Ripple correlation control is a nonlinear control approach applicable to power electronic circuits. It makes use of voltage, current, or power ripple and correlates this with switching functions to affect control [22]. The RCC solves major problems through less complex implementation. The most important factors and advantages of the RCC are the simple circuit implementation, fast computation/simulation time, there is no need for external perturbation like in P&O and IC, to generate ripple contents, converges asymptotically to the object and its converging rate can be tuned by the controller gain [21]. RCC is a method that is used to calculate the duty cycle which provides maximum power, which will be supplied to the gate of switching circuits to maintain MPP. The main advantage of RCC is that it uses inherent ripples that occur due to the power electronic elements i.e., the DC-DC convertor used in the PV system. Through the correlation of the time based derivative of voltage and power the RCC tries to identify whether this correlation is greater than zero i.e., to the left of the MPP, or less than zero i.e., to the right of the MPP, or exactly zero i.e., equal to MPP.

$$\frac{dp_{PV}}{dt}\frac{dv_{PV}}{dt} > 0\ when\ V_{PV} < V_M, \tag{4}$$

$$\frac{dp_{PV}}{dt}\frac{dv_{PV}}{dt} < 0\ when\ V_{PV} > V_M, \tag{5}$$

$$\frac{dp_{PV}}{dt}\frac{dv_{PV}}{dt} = 0 \; when \; V_{PV} = V_M. \quad (6)$$

The general equation used to find the ripple content is as:

$$\tilde{x}(t) = x(t) - \bar{x}(t) \quad (7)$$

$x(t)$ is the general quantity that can be array current, voltage or power which contains both the ripple as well as moving average component. $\tilde{x}(t)$ is the ripple content whereas $\bar{x}(t)$ is the average component. From equation (7) we can easily find out the ripple content of voltage, current and power.

Figure 3 represents the RCC that has been implemented to find out the duty cycle. This RCC is different from [21-22], as it has low pass filter instead of high pass filters as in [21-22]. In this work RCC mentioned in [25] is implemented for battery operated systems for output power of 95 watts. The work differs from [25], in terms of load which is battery in this work and converter type, which is buck converter, instead of boost converter used in for grid connected load as in [25]. From the Figure 3 error is equal to the product of the ripple content of both the power and voltage.

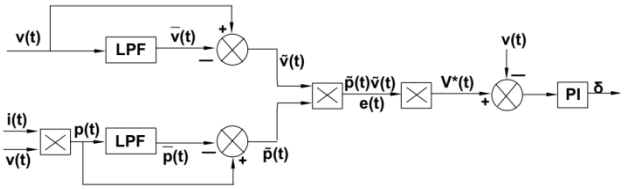


Fig. 3. RCC block of the MPPT Controller

To obtain the error $e(t)$ of power we will use the following mathematical equations;

$$p(t) = v(t) \times i(t) \quad (8)$$

Expressing i(t) and v(t) in terms of their ripple content as in (7) and using them in (8) we get,

$$p(t) = \bar{v}(t) \times \bar{\imath}(t) + \bar{\imath}(t) \times \tilde{v}(t) + \bar{v}(t) \times \tilde{\imath}(t) + \tilde{v}(t) \times \tilde{\imath}(t) \quad (9)$$

Therefore, power ripple can be expressed as;

$$\tilde{p}(t) = \bar{\imath}(t) \times \tilde{v}(t) + \bar{v}(t) \times \tilde{\imath}(t) + \tilde{v}(t) \times \tilde{\imath}(t) \quad (10)$$

We can further express the product of $\tilde{p}(t)$ and $\tilde{v}(t)$ as follows;

$$\tilde{p}(t) \times \tilde{v}(t) = \tilde{v}^2(t)\left[\bar{\imath}(t) + \bar{v}(t)\frac{\tilde{\imath}(t)}{\tilde{v}(t)}\right] + \tilde{v}^2(t)\tilde{\imath}(t) \quad (11)$$

Upon taking the derivative of $p(t)$in equation (21) w.r.t. $v(t)$ we get;

$$\frac{dp(t)}{dv(t)} = i(t) + v(t)\frac{di(t)}{dv(t)} \quad (12)$$

Linearizing, Figure 2 at point v_0, i_0, we get;

$$\left(\frac{di(t)}{dv(t)}\right)_{v_0} = \frac{\tilde{\imath}(t)}{\tilde{v}(t)} \quad (13)$$

Using equation (12) and (13) error can be expressed as;

$$\tilde{p}(t) \times \tilde{v}(t) = \tilde{v}^2(t)\left[\frac{dp(t)}{dv(t)}\right] + \tilde{v}^2(t)\tilde{\imath}(t) \quad (14)$$

Error is directly proportional to the magnitude of dp/dv as the average value $\tilde{v}^2(t)\tilde{\imath}(t)$ is zero over a cycle, $e(t)$ represents the distance from the MPP. When the output is on the left of MPP the average value of error is positive. When output is on the right of MPP the average value of error is negative. When the output is at the MPP the average value of error is zero. From Figure 3, it can be inferred that when the error signal passes through the first PI controller it generated the reference signal, and the difference of this reference signal & PV array voltage $v(t)$, acts as an input for the second PI controller that gives us the load angle δ, which is used to generate the PWM signal to control the switching of buck convertor.

The DC-DC converter is used to convert one level of DC voltage into another level. According to the requirements one can opt Boost, Buck or Buck-boost converter. In this work a DC-DC buck converter has been used [26]. The primary function of the buck converter is to decrease the voltage according to battery charging requirements with increase in the current for fast charging of battery.

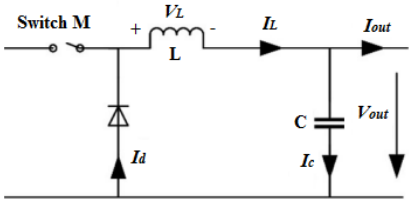


Fig. 4. Electrical Model of DC-DC Buck Converter.

III. SIMULINK MODEL OF PV SYSTEM WITH RCC MPPT

The Simulink model schematic of the PV system that is used for simulations is depicted in Figure 4, and represents a PV solar panel connected to a resistive load through a dc-dc buck converter with a variant subsystem of MPPT controller that allow to choose between these three MPPT algorithms. P&O, incremental conductance and RCC.

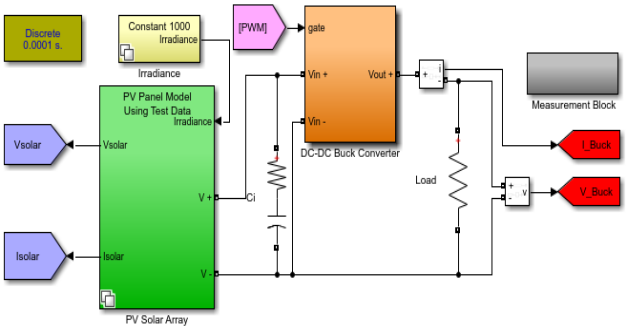


Fig. 5. Simulink Model of PV System.

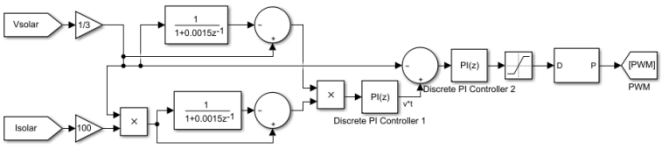


Fig. 6. MPPT variant blocks and Simulink implementation of RCC MPPT controller [25].

The PV parameters used in Simulink simulation model of this work are given in Table I. The components parameters of the buck converter used in Simulink are given in Table II. The components parameters of RCC block are given in Table III.

TABLE I
PV ARRAY SPECIFICATIONS

S.No.	Parameter Name	Value
1	Number of cells in series	n-cells = 36
2	Open circuit voltage	V_{oc}= 21.6 V
3	Short circuit current	I_{sc}= 7.34 A
4	Series resistance of PV model	R_s= 0 Ω
5	Parallel resistance of PV model	R_p= ∞ Ω
6	Diode quality factor of PV model	N=1.5
7	DC link capacitor	C_{dc}= 100 μF

TABLE II
BUCK CONVERTER COMPONENTS PARAMETERS

S.No.	Parameter Name	Value
1	Inductance	$L = 23\ \mu H$
2	Capacitance	$C = 120\ \mu F$
3	Switching Frequency	$f_{sw}=2*10^6$
4	Sample Time	0.0001 SEC
5	Diode Resistance	$R_{ON} = 0.001\Omega$

TABLE III
RCC COMPONENTS PARAMETERS

S.No.	Parameter Name	Value
1	Time constant of LPF's	1.5 ms
2	PI controller (inside MPPT block)	K_p= 200
		K_I= 5.5
3	PI controller (outside MPPT block)	K_p= 2e-9
		K_I= -0.009

Simulation has been performed for three different types of variations, fist one is at 1000 W/m^2 at 25^{o}C, second at 800 W/m^2 at 25^{o}C and third is for step variation [5] as shown in Figure 7.

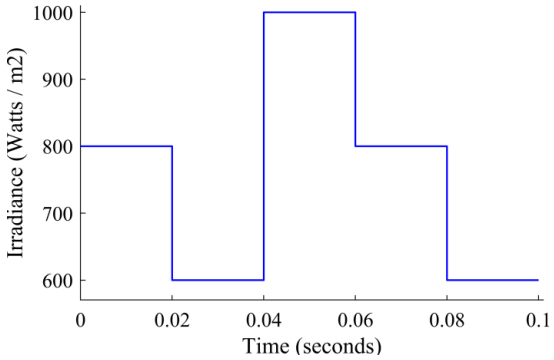


Fig. 7. Step Variation of irradiance [5].

IV. RESULTS AND DISCUSSIONS

All the simulations have been performed in MATLAB/SIMULINK 2018. First the irradiation level is fixed at 1000 W/m^2 at 25^{o}C, and results of voltage, current and power for all three methods i.e. perturb & observe, incremental conductance and ripple correlation control have been captured using measurement scopes. Then same process is performed for 800 W/m^2 irradiance level at 25^{o}C. After this simulation is performed for Step variation of irradiance as shown in Figure 7, and graph were captured. There are many research papers available of P&O and IC methods in which simulation results have been given. In this work PV model parameters have been selected as in [5], since it was also a battery connected system.

Figures 8, 9, and 10 shows the simulation output of P&O MPPT obtained by the simulation using the setup mentioned in figure 5. The simulation results shows that conventional P&O MPPT is very much affected from being slow tracking and oscillations both. This is because of internal ripples, which are generated due to the use of power electronics components. Actually, the system in P&O case, never achieved stability, and it always oscillate around MPP point. It is due to the internally generated ripples, which are considered as perturbations in P&O MPPT. Hence, the performance of conventional P&O MPPT is not up to the mark in terms of stability, ripples and tracking.

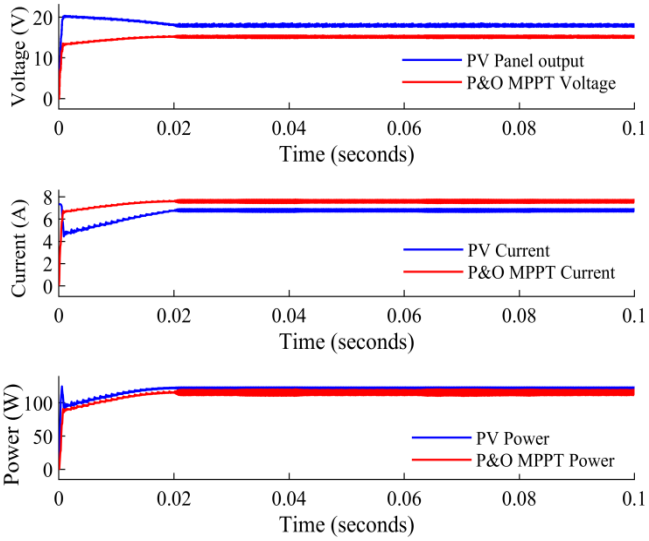


Fig 8. Voltage, Current and Power output of P&O MPPT for 1000 W/m^2 irradiance level at 25^{o}C.

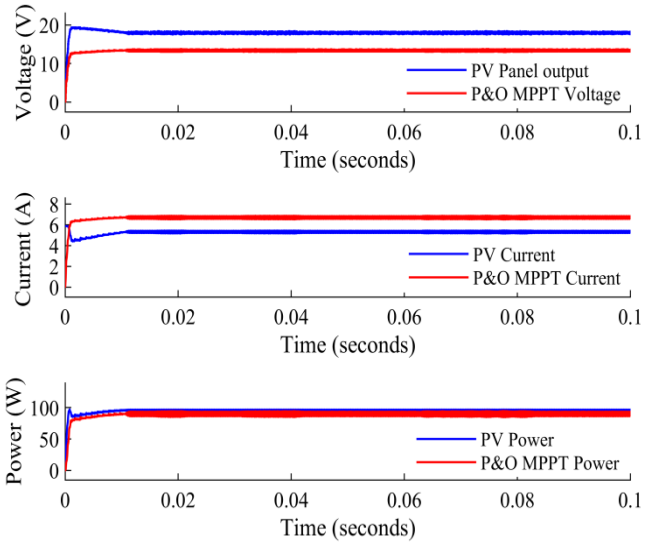


Fig 9. Voltage, Current and Power output of P&O MPPT for 800 W/m^2 irradiance level at 25^{o}C.

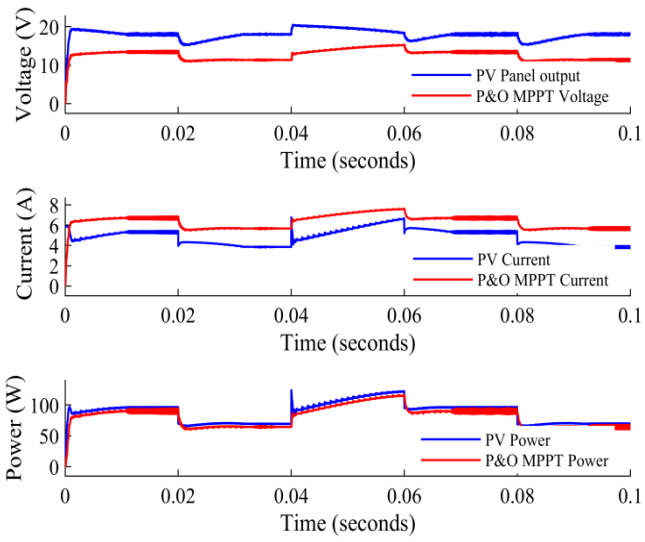


Fig 10. Voltage, Current and Power output of P&O MPPT for Step irradiance at 25^{o}C.

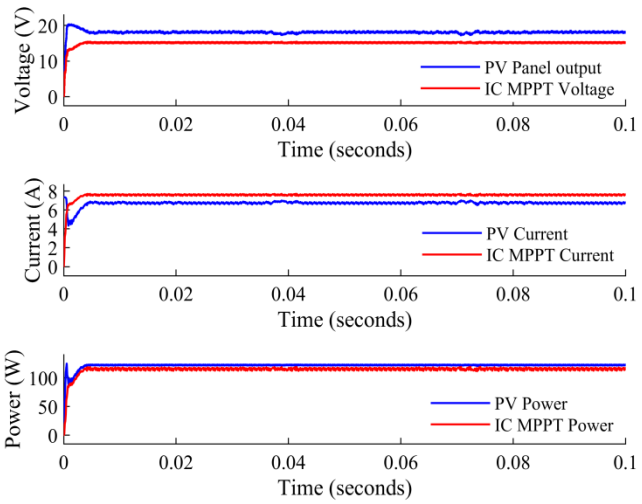


Fig 11. Voltage, Current and Power output of IC MPPT for 1000 W/m^2 irradiance level at 25^{o}C.

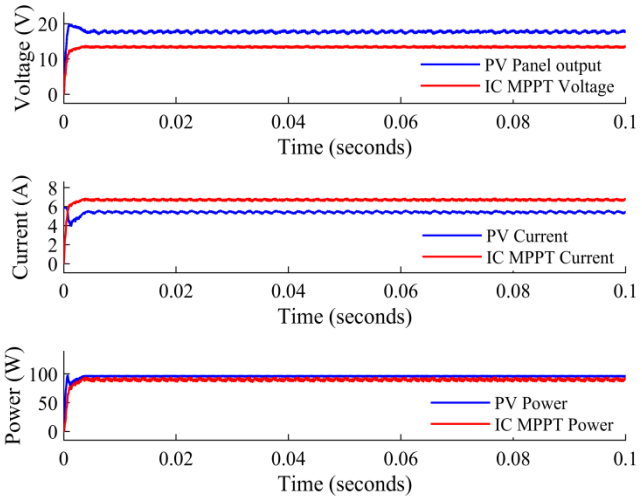


Fig 12. Voltage, Current and Power output of IC MPPT for 800 W/m^2 irradiance level at 25^{o}C.

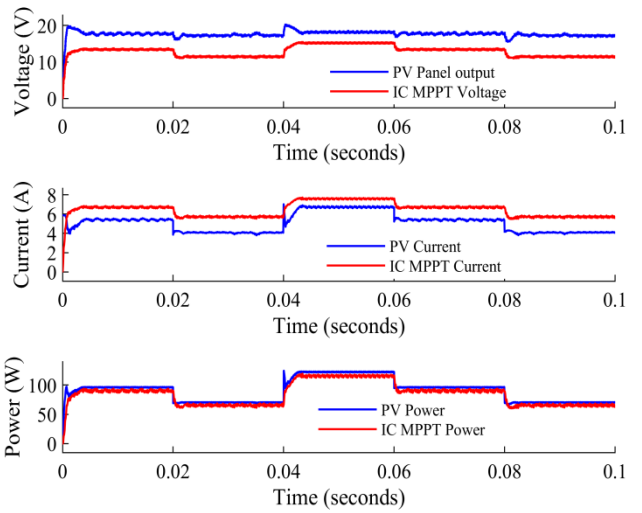


Fig 13. Voltage, Current and Power output of IC MPPT for Step irradiance at 25^{o}C.

Figures 11, 12 and 13 shows the simulation output of IC MPPT, from these Figures it can be clearly seen that, IC tracking is better than P&O MPPT, but output still consists of oscillations in all three cases of irradiance.

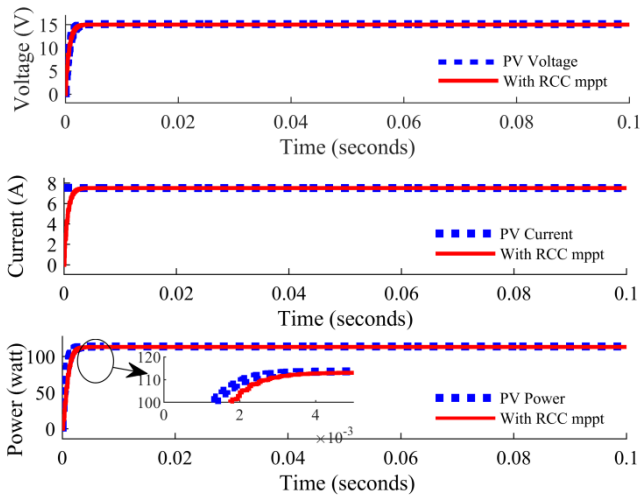


Fig 14. Voltage, Current and Power output of RCC MPPT for 1000 W/m^2 irradiance level at 25^{o}C.

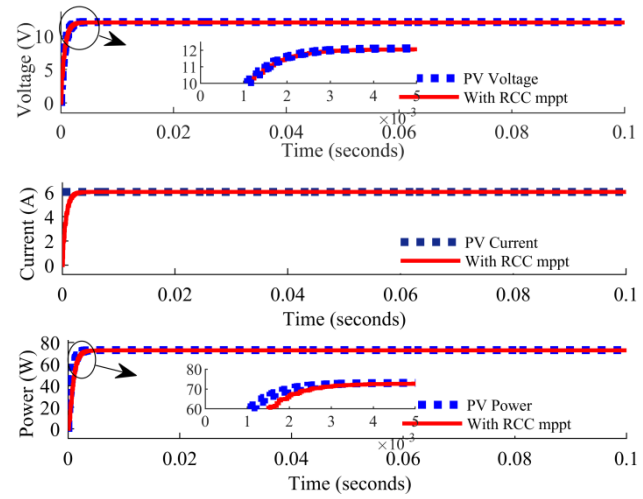


Fig 15. Voltage, Current and Power output of RCC MPPT for 800 W/m^2 irradiance level at 25^{o}C.

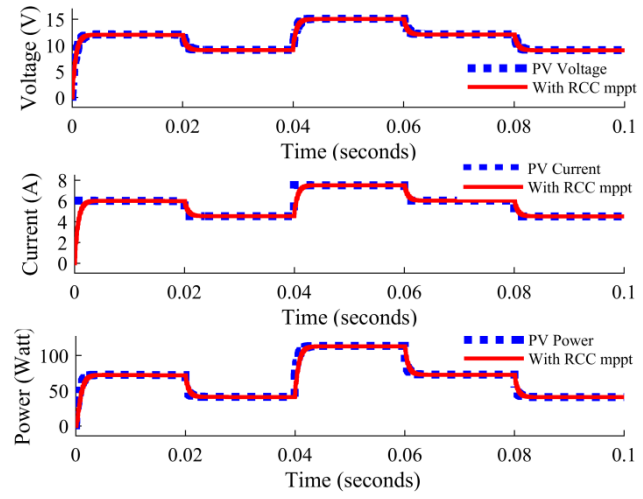


Fig 16. Voltage, Current and Power output of RCC MPPT for Step irradiance at 25^{o}C.

From the Figures 14-16, it can be clearly seen that the problem of slow tracking and ripples (oscillations) has been solved by RCC MPPT. Simulation results verify that RCC MPPT is capable of solving both these problems effectively. Basically these ripples were internal, generated due to use of power electronic converters, and RCC utilizes these internal

ripples as perturbations, and is able to eliminate them. Ripple problem is completely solved by RCC MPPT, but for tracking problem can be further improved by some adaptive control strategies.

V. CONCLUSION

This paper presents a comparative study between perturb & observe, incremental conductance and ripple correlation control MPPT. The PV system model used for simulation consists of the PV panel, the variant subsystem of irradiance, the buck converter and the variant subsystem of MPPT controller. This comparative analysis aims to show the improved performance of RCC method over conventional MPPT methods like P&O and IC. The simulation results show that both the methods perturb & observe and incremental conductance have oscillations in current output and hence power output in all three case i.e. 1000 w/m^2, 800 w/m^2 and step irradiance levels; whereas, RCC MPPT does not has any sort of ripples neither in current output, nor in power output, which is a good sign for longer battery life.

REFERENCES

[1] Femia Nicola, Petrone Giovanni, Spagnuolo Giovanni. Optimization of perturb and observe maximum power point tracking method. *IEEE Trans., on Power Electronics* 2005; 20(4). 963–73.

[2] Pandey Ashish, Dasgupta Nivedita, Mukerjee Ashok Kumar. High-performance algorithms for drift avoidance and fast tracking in solar mppt system, *IEEE Trans., on Energy Converters* 2008;23. 681–9.

[3] Abdelsalam Ahmed K, Massoud Ahmed M, Ahmed Shehab, Enjeti Prasad N., High-performance adaptive perturb and observe MPPT technique for photovoltaic-based micro grids. *IEEE Trans., on Power Electronics* 2011; 26(4). 1010–21.

[4] Mohammed A, Elgendy Bashar, Zahawi, Atkinson David J. Assessment of perturb and observe MPPT algorithm implementation techniques for PV pumping applications. *IEEE Trans., on Sustainable Energy* 2012; 3(1). 21–33.

[5] Ioan Viorel Banu et al., "Comparative analysis of the perturb-and-observe and incremental conductance MPPT methods", 8TH *IEEE International Symposium on Advanced Topics in Electrical Engineering (ATEE),* 2013.

[6] Zainuri Muhammad Ammirrul Atiqi Mohd, Radzi Mohd Amran Mohd, Soh Azura Che, Rahim Nasrudin Abd. Development of adaptive perturb and observe-fuzzy control maximum power point tracking for photovoltaic boost dc–dc converter. *IET Renew Power Gener*., 2014;8(2). 183–94.

[7] Kollimalla Sathish Kumar, Mishra Mahesh Kumar. A novel adaptive P&O MPPT algorithm considering sudden changes in the irradiance. *IEEE Trans., on Energy Converters*, 2014; 29(3). 602–10.

[8] Mohammed A Elgendy, Zahawi Bashar, Atkinson David J. Operating characteristics of the P&O algorithm at high perturbation frequencies for standalone PV systems. *IEEE Trans., on Energy Converters*, 2015; 30(1). 189–98.

[9] Ahmed Jubaer, Salam Zainal. "An improved perturb and observe (P&O) maximum power point tracking (MPPT) algorithm for higher efficiency", *Applied Energy* 2015;150. 97–108.

[10] Raedani Ronn Hanif Moin, "Design, testing and comparison of P&O, IC and VSSIR MPPT technique," *In proceedings of the 3rd international conference on renewable energy research and applications*, 2014, p. 19–22.

[11] LeeJae., HoBaeHyunSu., and ChoBo Hyung., "Advanced incremental conductance MPPT algorithm with a variable step size.", *In proceedings of the 12th International power electronics and motion control conference*, 2006. EPEPEMC 2006.

[12] Li Jiyong, Wang Honghua. A novel stand-alone PV generation system based on variable step size INC MPPT and SVPWM control. *IEEE Conf.,PEMC* 2009. 2155–60.

[13] Hsieh Guan-Chyun, Hsieh Hung-I, Tsai Cheng-Yuan, Wang Chi-Hao. Photovoltaic power-increment-aided incremental-conductance MPPT with two phased tracking. *IEEE Trans Power Electr.,* 2013;28(6). 2895–911.

[14] Uma Maheswara Bala Murali, Rao Ch V.Krishna, Srihari Babu A, Suman S. Design and simulation of PV system with incremental conductance method for maximum power point tracking. *International Journal of Scientific Engineering and Technology* (ISSN. 2277-1581); 3(5). 2014. p. 643-46.

[15] Sivakumar P, Kader Abdullah Abdul, Kaliavaradhan Yogeshraj, Arutchelvi M. Analysis and enhancement of PV efficiency with incremental conductance MPPT technique under non-linear loading conditions, *Renew., Energy* 2015; 81. 543–50.

[16] Veerachary M, Senjyu T, Uezato K. Maximum power point tracking control of IDB converter supplied PV system. *IEEE Proceeding on Electronics Power Applications* 2001. 494–502.

[17] Weidong X, Dunford WG. A modified adaptive Hill Climbing MPPT method for photovoltaic power systems. *IEEE Power Electron Spec Conference*, 2004;35. 1957–63.

[18] Liu Fangrui, Kang Yong, Zhang Yu and Duan Shanxu. Comparison of P&O and Hill Climbing MPPT Methods for Grid-Connected PV Converter. *In Proceedings of the 3rd IEEE conference on industrial electronics and applications*, ICIEA 2008.

[19] J. Prasanth Ram, T. Sudhakar Babu, N. Rajasekar, Renewable and Sustainable Energy Reviews, *Renewable and Sustainable Energy Reviews* 67 (2017) 826–847.

[20] Mingxuan Mao et al., "Maximum Power Point Tracking for Cascaded PV-Converter Modules Using Two-Stage Particle Swarm Optimization", *Nature Scientific Reports*, Article number. 9381, 2017.

[21] Pallab Midya et al., "Dynamic Maximum Power Point Tracker for Photovoltaic Applications", *IEEE Power Electronics Specialists Conference*, 1996.

[22] Philip T.Krein, "Ripple Correlation Control, With Some Applications", *IEEE International Symposium on Circuits and Systems (ISCAS),* 1999.

[23] Raghav Khanna et al., "Maximum Power Point Tracking Using Model Reference Adaptive Control", *IEEE Trans., on Power Electronics*, 2014.

[24] Alessandro Costabeber et al., "Convergence Analysis and Tuning of a Sliding-Mode Ripple-Correlation MPPT", *IEEE Trans., on Energy Conversion*, 2014.

[25] Satish R et al., "A Maximum Power Point Tracking Technique Based on Ripple Correlation Control for Single-Phase Single-Stage Grid Connected Photovoltaic System", Science Direct, *Energy Procedia*, 2016.

[26] Salman Salman, Xin AI* and Zhouyang WU, Design of a P-&-O algorithm based MPPT charge controller for a stand-alone 200W PV system, Springer Protection and Control of Modern Power Systems, 2018.

Computational model for temperature in tree trunk for energy harvesting

Yajun An
School of Interdisciplinary Arts & Sciences
University of Washington, Tacoma
Tacoma WA, USA
yajuna@uw.edu

Michael Hockman
School of Engineering and Technology
University of Washington, Tacoma
Tacoma WA, USA
hockmm@uw.edu

Orlando Baiocchi
School of Engineering and Technology
University of Washington, Tacoma
Tacoma WA, USA
baiocchi@uw.edu

Selina Teng
College of Engineering
University of Washington, Seattle
Seattle WA, USA
seliteng@uw.edu

***Abstract*—The temperature difference in the cross section of a tree trunk, measured radially, generates voltage. This is leveraged to power thermoelectric devices for the purpose of environmental monitoring. The authors present a numerical model for the temperature differential within a tree stem. This model contributes to a proof of concept for the feasibility of using trees as an ambient power source for wireless devices.**

The University of Washington Tacoma is in ongoing research cooperation with the University of Paraiba, Brazil and the University of the Azores, Portugal in the areas of Wireless Sensor Networks, Energy Harvesting, and Environmental Monitoring. This paper is part of these collaborative efforts and is related to the specific goal of deploying a network of wireless environmental sensors in the Island of Sao Miguel, Archipelago of the Azores, to monitor both the natural (volcanic) and industrial pollution, as well as monitoring the traffic of tourists on the island.

***Index Terms*—Heat Equation, Thermoelectric Generator (TEG), Computational Solutions, Numerical Model, Trees, Energy Harvesting**

I. Introduction

UNESCO, as part of its Geopark program, protects the Archipelago of the Azores, where the constant influx of tourists in recent years endangers the ecosystem through the pollution of the environment and the damage of protected areas. Numerical data is needed to inform policy makers about the severity of the environmental impact of tourism. Hence, the authors and collaborators at the University of the Azores (whose campus is located on the Island of Sao Miguel in the Archipelago of the Azores) propose a wireless sensor network to perform environmental monitoring in Sao Miguel. In the current phase of the project, we built a similar experiment in the Pierce County area in Washington State. The considerations are fundamentally the same, but the latitude in Washington requires us to investigate temperature distribution in tree trunks along difference directions, whereas directions can be treated as the same in the equatorial Azores.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

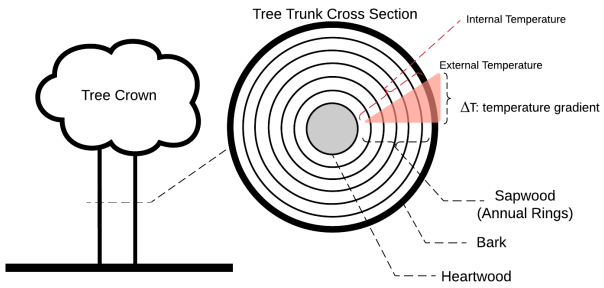


Fig. 1: Tree Stem Cross Section. Source: [2]

The deployment of sensor networks in remote natural areas is a challenge regarding the energy requirements of the sensors themselves and the transmission systems. Since the use of batteries is both inconvenient and environmentally unfriendly, alternative processes have been considered. In particular, energy harvesting from trees is a potential candidate.

Healthy trees maintain a temperature of approximately 21.4 degrees Celcius in the leaves as a precondition for photosynthesis [1]. In order to thermoregulate, the tree stores a substantial amount of heat in its stem or trunk. The annual rings in a tree trunk (Fig. 1) can thus be treated as isothermal subvolumes. Reference [2] asserts that the temperature gradient between any annual ring and the external temperature is either approximately constant or varies with some time-delay as the external temperature varies.

In order to utilize the small amount of energy obtained from the trees, it is important to know the heat distribution inside the trees as a function of the nature of the trees, the solar radiation, and the specific characteristics of the environment in which the harvesting takes place. That is the purpose of this paper, which uses mathematical modeling and simulation of the conversion of heat into electricity via the Seebeck effect.

Fig. 2: TEG devices installed on tree trunks in Washington State

To harvest the energy, we use a thermoelectric generator attached to a metal rod, which is driven laterally through the tree stem. Although they are less efficient than heat engines, thermoelectric generators (TEGs) are preferred in remote applications that only require low power, which suits the relatively low output expected from natural heat sources such as trees [2]. The use of tree stems as a heat source for TEGs is promising because it is a long-lasting, maintenance-free energy source, suitable for powering local wireless devices, such as environmental sensors. Our group hypothesizes that the most crucial element of the TEG system is the way in which the temperature of the interior of the tree is brought into contact with the TEG.

In this work, we present mathematical models simulating the temperature distribution in tree trunks with partial differential equations (PDEs) and their numerical solutions, with the purpose of examining the efficacy of trees for energy harvesting. [3] presented an earlier 2D model to simulate tree temperature distributions, driven by the influences of solar radiation, infrared emission and absorption, convection, and conduction. Their numerical solutions were obtained by a Finite Difference (FD) scheme, with centered differencing in space, and leapfrog in time. Another 2D temperature distribution model for tree trunks, FireStem2D, has been created by [4]. FireStem2D has parameters for studying the effects of forest fires on tree stem injury, whereas the present paper examines the temperature of trees in ambient conditions.

The present work focuses on a simplified 1D PDE model that retains the needed sensitivity for the energy harvesting task. In the next section, we justify the simplification obtained by reducing the dimensions from 3D to 2D, then to 1D). We follow the source terms used in Potter and Andresen model [3]. In handling heat source, we treat the heat flux as the only source and solve a boundary value problem from the converted temperature condition. We apply a homogenous Neumann condition at the center of the tree trunk [4], and a nonhomogenous Dirichlet condition at the bark surface. We solve the differential equation in one dimension numerically using the Crank-Nicolson method, which is superior to explicit methods in its stability, thereby reducing the computation cost.

This project is a work in progress; we present PDE models and their numerical solutions based on example data from current literature [3], since the COVID-19 pandemic has impacted our plans to collect field data in Washington State. Our result should provide reference for hardware analysis, particularly, giving insight to the magnitude of voltage possible for the thermoelectric devices. First we describe the PDE, including the justification for dimension reduction, source terms, and example input data. We describe the implementation of the FD scheme in polar coordinates with the Crank-Nicolson method, and boundary conditions for the core and bark surface of the tree stem. Finally, we discuss the results of our numerical experiments and limitations of our current model, and describe ongoing and future steps.

II. PDE MODELING FOR TEMPERATURE DISTRIBUTION IN TREE TRUNKS

Heat transfers within the tree trunk vertically, radially, and tree temperature might differ at different azimuth angle. To fully model the temperature distribution within a tree trunk, we need to study the heat equation in three dimensions.

$$\begin{aligned}\rho c\frac{\partial T}{\partial t} &= \nabla\cdot(k\nabla T)\\ &= \frac{1}{r}\frac{\partial}{\partial r}\left(kr\frac{\partial T}{\partial r}\right)+\frac{1}{r}\frac{\partial}{\partial \phi}\left(\frac{k}{r}\frac{\partial T}{\partial \phi}\right)+\frac{\partial}{\partial z}\left(k\frac{\partial T}{\partial z}\right)\end{aligned}\tag{1}$$

Here T represents temperature (K), ρ density (kg/m^3), c specific heat (J/(kgK)), t time (s), k thermal conductivity (W/(mK)), r distance from center of the tree (m), and ϕ azimuth angle, measured clockwise with south being $\phi = 0$ (radian/degree), we assume $\partial k/\partial\phi = 0$ [3]. For our case we use the following values of thermal conductivity k and heat capacity ρc that were found in [5]:

- Thermal Conductivity: $k = 0.12\frac{W}{mK}$
- Heat Capacity: $\rho c = 1.7\frac{J}{KgK}$

For our application however, we can greatly simplify the equation with reasonable assumptions. As our devices would be installed in forests with trees of similar age and growth, it is reasonable to assume that we install all devices at the same height - this eliminates the dependence on the height (z-variable).

Therefore we consider a two dimensional simplification to the above model.

A. Simplification of temperature distribution model

We eliminate the dependence on height z and obtain the following standard heat equation for temperature distribution in two dimensional cylindrical coordinates:

$$\begin{aligned}\rho c\frac{\partial T}{\partial t} &= \nabla\cdot(k\nabla T)\\ &= \frac{1}{r}\frac{\partial}{\partial r}\left(kr\frac{\partial T}{\partial r}\right)+\frac{1}{r}\frac{\partial}{\partial \phi}\left(\frac{k}{r}\frac{\partial T}{\partial \phi}\right)\end{aligned}\tag{2}$$

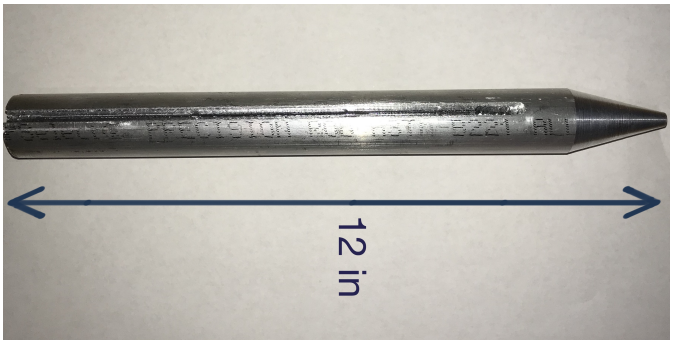


Fig. 3: Aluminum rod, to be inserted in tree trunk horizontally along a radius.

In the above simplified model, heat distribution depends on r, as well as ϕ. For our application, we primarily care about the temperature differences during different time of the day along one direction (see Fig. 3), and install all devices in the direction where such differences occur at the largest magnitude. Hence, we seek to model heat distribution only in the directions of largest temperature difference.

As seen in Fig. 2 of [3], in the northern hemisphere, the largest temperature differences occur in south and west directions. However, due to the limited access to data, we can only investigate the south and north aspect in this work. This further reduces the heat equation to a one dimensional problem and eliminates the ϕ dependence. Finally, we obtain a one-dimensional temperature PDE model.

$$\begin{aligned}\rho c\frac{\partial T}{\partial t} &= \frac{1}{r}\frac{\partial}{\partial r}\left(kr\frac{\partial T}{\partial r}\right)\\ &= k\frac{1}{r}\frac{\partial T}{\partial r}+k\frac{\partial^2 T}{\partial r^2}.\end{aligned} \tag{3}$$

III. Boundary conditions, source term and data

The inner boundary condition is a no-flux condition

$$\left.\frac{\partial T}{\partial r}\right|_{r=0}=0, \tag{4}$$

as justified by [4]. We use a Dirichlet condition at the outer boundary, and approximate the surface temperature by recorded ambient temperature from Tacoma during a 12 hour cycle starting at 7 am.

The source term is made up from several types of sources, including convective heat loss/gain when the tree surface temp is different from the ambient temp (both free and forced), direct and diffusion solar radiation, and wave radiation from and to the tree as black bodies.

A. Convective heat loss/gain

We denote the heat loss/gain by H [3], and

$$H = h(T_{sfc}-T_{air}), \tag{5}$$

with h (W/(m^2K)) being the convective heat transfer coefficient. Here

$$h = h_{free}-h_{forced}, \tag{6}$$

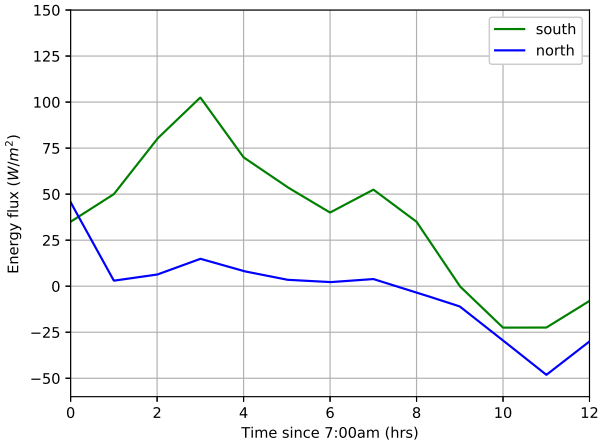


Fig. 4: Linear combination of the source terms, $H+(1-\alpha)(S_{dir}+S_{dif})+(IR_{in}-IR_{out})$ on the north and south sides of the tree. Source: [3]

B. Solar radiation

$S_{dir}+S_{dif}$ represent direct solar radiation, plus diffusion solar radiation.

$$S_{dir} = S_0(\tau)^{\sec(Z)}\cos(i) \tag{7}$$

and

$$S_{dif} = \frac{S_0\cos(Z)}{3}(1+\cos(Z))(\eta-\tau^{\sec(Z)}) \tag{8}$$

Where Z is the solar zenith angle (angle between the sunlight and the normal of the flat ground), and i is the angle of incidence (angle between the direction of the sunlight and the normal of the tree surface), S_0 = 1368 $\frac{W}{m^2}$ τ is 0.76 to 0.81, and η is 0.80 to 0.84 [3].

The amount of solar radiation the tree receives also depends on the albedo, α, which is the amount of solar radiation reflected off the surface of the tree. In the end, the term we see as a source term is

$$(1-\alpha)(S_{dif}+S_{dir}) \tag{9}$$

C. Net Infrared

The net infrared $IR_{in}-IR_{out}$ is the final source term and represents the difference between the radiation going to and the radiation coming from the tree.

The net infrared depends on σ, the Stefan-Boltzmann constant $5.67\times10^{-8}\frac{W}{m^2K^2}$ as well as the temperature of the surface of the tree and the temperature of the surrounding air. The exact relationship is as follows

$$IR_{out} = \sigma T_{sfc}^4 \tag{10}$$

and

$$IR_{in} = \sigma T_{air}^4 \tag{11}$$

D. Explanation of example data

Figure 4 shows the linear combination of the source terms $H + (1-\alpha)(S_{dir} + S_{dif}) + (IR_{in} - IR_{out})$ on the north and south sides of the tree from [3]. Due to the COVID-19 pandemic, our hardware team was not able to collect these data in the local forest. We therefore use the aspen simulations data found in [3]. One of the future aspect of this work is to collect experimental data locally, and run our numerical simulation again.

IV. Finite difference (FD) schemes for the heat equations

We compute the temperature distribution in tree trunks numerically with the Finite difference method, as it is simple to implement, and robust [6]. Here in particular, we discretize the continuous equation in one dimension using the Crank-Nicolson scheme, due to its advantage in stability, and higher accuracy in time [7]. The radius of the tree stem is approximately 12 inches between the exterior and core in moderately sized trees we target for our study. The aluminum rod (Fig.3) in our TEG hardware system is about 12 inches in length with a 0.5 inch radius, embedded in the trunk of a tree at approximately breast height. Due to the size comparisons of our devices and the tree trunk, it is more appropriate to model the heat equation in polar form (3) than the standard Cartesian form. Let $T_i^n \approx T(r_i, t_n)$ denote the numerical solution at the point (r_i, t_n), Δt the time step, and Δr grid size. With second order central difference in space

$$\frac{\partial T}{\partial r} \approx \frac{T_{i+1} - T_{i-1}}{2\Delta r} \tag{12}$$

$$\frac{\partial^2 T}{\partial r^2} \approx \frac{T_{i-1} - 2T_i + T_{i+1}}{(\Delta r)^2} \tag{13}$$

and first order forward in time

$$\frac{\partial T}{\partial t} \approx \frac{T^1 - T^0}{\Delta t}, \tag{14}$$

we obtain the Crank-Nicolson method by averaging the above spatial differences in time:

$$\begin{aligned}(T_i^{n+1} - T_i^n)/\Delta t = 1/(4ar_i\Delta r)(T_{i+1}^n - T_{i-1}^n + T_{i+1}^{n+1} - T_{i-1}^{n+1}) \\ + 1/(2a(\Delta r)^2)(T_{i-1}^n - 2T_i^n + T_{i+1}^n + T_{i-1}^{n+1} - 2T_i^{n+1} + T_{i+1}^{n+1}),\end{aligned} \tag{15}$$

where $a = \frac{\rho c}{k}$.

We implement boundary conditions in the numerical scheme. For the boundary at the center of the tree, we follow the same no-flux Neumann condition in [4]. We use the second order central difference approximation at $r = 0$:

$$\left.\frac{\partial T}{\partial r}\right|_{r=0} \approx \frac{T_2 - T_0}{2\Delta r} = 0. \tag{16}$$

Equation (15) gives a representation for the ghost cell value T_0 in terms of T_2, which is implemented in the first row of the matrix below.

For the outer boundary condition at the surface of the tree, we implement the Dirichlet condition with ambient temperature $g_1(t)$

$$\left.T(r,t)\right|_{surface} = g_1(t), \tag{17}$$

in the last row of the following matrix equation. The source term is also present in the last row of the following matrix equation.

Together, we solve the following system to compute the heat distribution in the tree trunk:

$$\begin{pmatrix} 1+\beta & -\beta & & 0 \\ \alpha_2 & 1+\beta & -\gamma_2 & \\ \ddots & \ddots & \ddots & \\ & \alpha_{m-1} & 1+\beta & -\gamma_{m-1} \\ 0 & & \alpha_m & 1+\beta \end{pmatrix} \begin{pmatrix} T_1^1 \\ T_2^1 \\ \vdots \\ T_{m-1}^1 \\ T_m^1 \end{pmatrix} =$$

$$\begin{pmatrix} (1-\beta)T_1^0 + \beta T_2^0 \\ -\alpha_2 T_1^0 + (1-\beta)T_2^0 + \gamma_2 T_3^0 \\ \vdots \\ -\alpha_{m-1}T_{m-2}^0 + (1-\beta)T_{m-1}^0 + \gamma_{m-1}T_m^0 \\ -\alpha_m T_{m-1}^0 + (1-\beta)T_m^0 + \gamma_m(g_1(t) + g_1(t+\Delta t)) + source \end{pmatrix}, \tag{18}$$

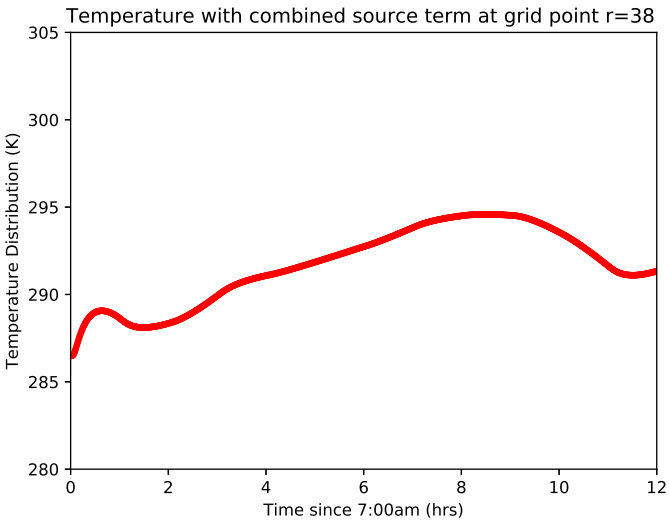
where $\alpha_i = \frac{\Delta t}{4ar_i\Delta t} - \frac{\Delta t}{2a(\Delta r)^2}$, $\beta = \frac{\Delta t}{a(\Delta r)^2}$, and $\gamma_i = \frac{\Delta t}{4ar_i\Delta t} + \frac{\Delta t}{2a(\Delta r)^2}$.

V. Numerical experiments and discussion of results

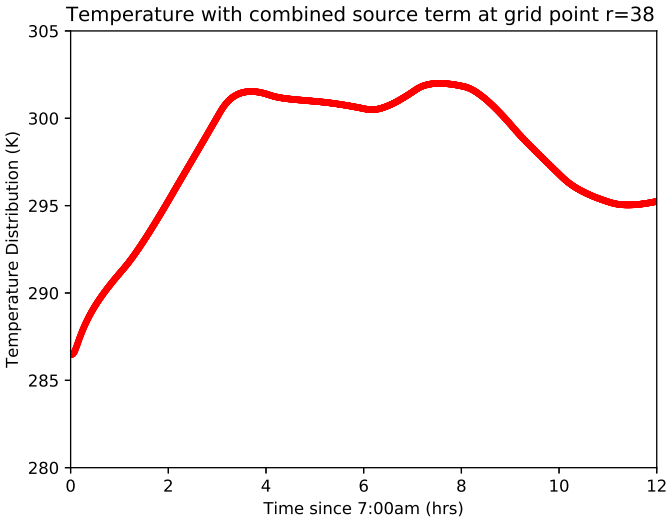
With the above FD scheme, combined example source term, and a constant temperature at 7 am as initial condition, we obtain the following figures of temperature distributions.

From Fig.5, Fig.6, and Fig.7, we conclude that the south aspects uniformly have a larger temperature range throughout the day; therefore it is preferable to install all TEG devices in the south. We also see that, the further away we are from the tree center, the trend of temperature is more consistent with heat flux. From our simulations, temperature distributions with medium depth don't differ significantly, for example, temperature measured at depth 7 to 9 inches have similar temperature ranges. We expected fairly constant temperature near the center of the trunk, but our simulation does not reflect that, and this is another issue we expect to address in our future work. With measurements of the tree surface temperature, our hardware team can adjust the length of the rod for optimal voltage generated from the temperature differences. The highest temperatures uniformly occur roughly 8 or 9 hours after we start the measurements, with lowest temperature in the beginning and end of the measurement cycle.

The simulation appears to be excessively warm overall for the north aspect; this is due to the limitation of the example data as described in [3], and our interpretation of the example data.

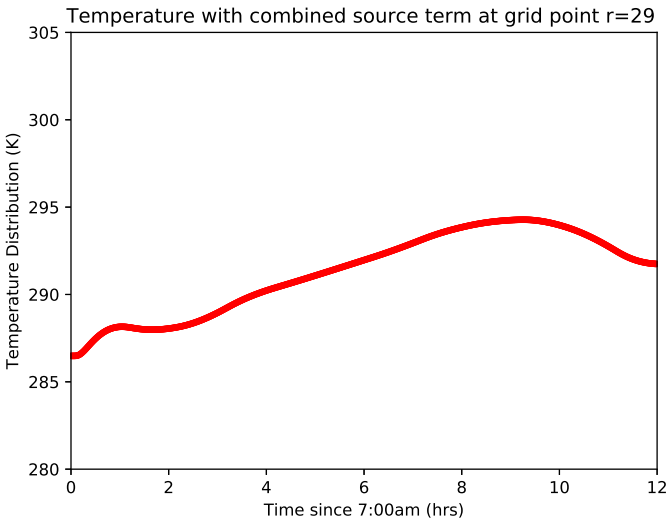


(a) Temperature distribution throughout the day, north aspect

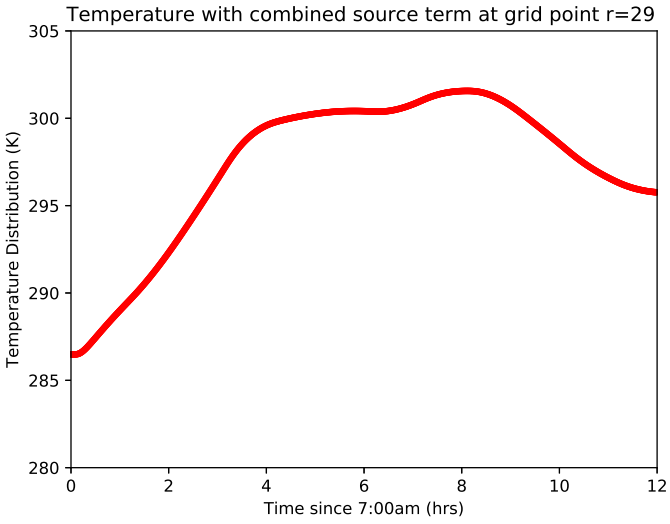


(b) Temperature distribution throughout the day, south aspect

Fig. 5: For a tree with 12 inch radius, temperature distribution in the north and south aspects, with depth of about 9 inches.



(a) Temperature distribution throughout the day, north aspect



(b) Temperature distribution throughout the day, south aspect

Fig. 6: For a tree with 12 inch radius, temperature distribution in the north and south aspects, with depth about 7 inches.

VI. Conclusion and on-going work

In this work, we modeled the temperature distribution in a tree trunk with the one dimensional heat equation in cylindrical coordinates, with various heat sources. We simplified the model to fit our purpose of studying the temperature distribution for applying TEG devices efficiently [8]. We applied a Crank-Nicolson FD scheme for numerical simulation, with example data, and analyzed the optimal aspect for installation of devices, as well as depth of installation.

We are working on higher dimension simulations for the temperature distribution for tree trunks, to gain more insight for other potential installation aspects. We will also be running these simulations locally once we have access to more data specific to tree species in western Washington State, as well as data from our collaborators in Brazil in Azores.

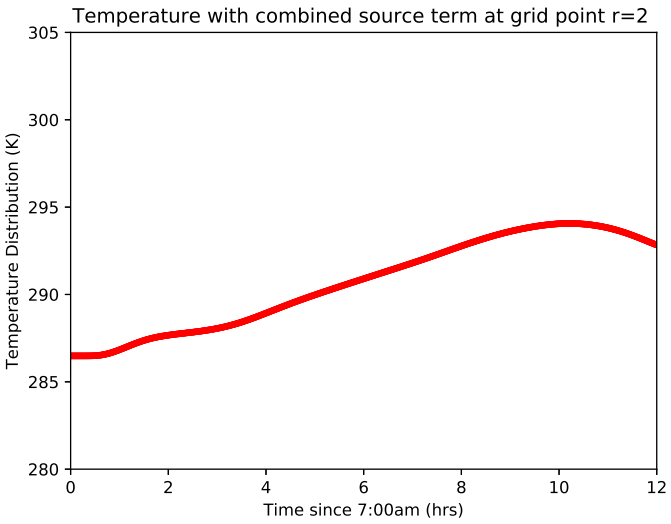
Acknowledgment

We appreciate the contribution to this work by Nick Purcell of the University of Washington (data collection), and Professor Cleonilson Protasio de Souza of the University of Paraiba.

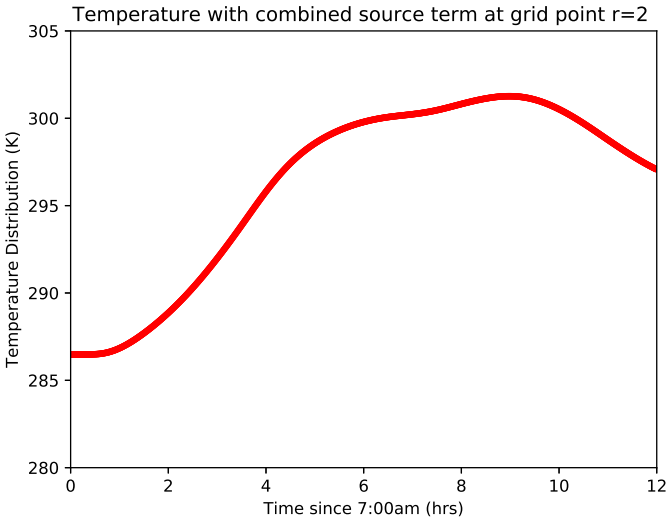
References

[1] B. R. Helliker and S. L. Richter, “Subtropical to boreal convergence of tree-leaf temperatures,” Nature, vol. 454, July 2008. [Online.] Available: Academic OneFile, http://find.galegroup.com. [Accessed July 21, 2020].

[2] P. Souza, F. Carvalho, F. Silva, H. Andrade, N. Silva, O. Baiocchi, and I. Müller, “On Harvesting Energy from Tree Trunks for Environmental Monitoring,” International Journal of Distributed Sensor Networks, pp. 1-9, 2016.



(a) Temperature distribution throughout the day, north aspect



(b) Temperature distribution throughout the day, south aspect

Fig. 7: For a tree with 12 inch radius, temperature distribution in the north and south aspects, near tree center.

[3] B. E. Potter and J. A. Andresen, "A Finite-Difference Model of Temperatures and Heat Flow within a Tree Stem," Canadian Journal of Forest Research, vol. 32, no. 3, pp. 548–555, 2002.

[4] E. K. Chatziefstratiou, et al., "FireStem2D – A Two-Dimensional Heat Transfer Model for Simulating Tree Stem Injury in Fires," PLoS ONE, vol. 8, no. 7, p.e70110, July 2013.

[5] W. Simpson, A. TenWolde, "Physical properties and moisture relations of wood," in *Wood handbook: wood as an engineering material*, Madison, WL: USDA Forest Service, Forest Products Laboratory, 1999, General technical report FPL, GTR-113: pp. 3.1-3.24

[6] S. Linge, H. P. Langtangen, "Diffusion Equations," in *Finite Difference Computing with PDEs*, vol. 6, 1st ed. Springer, 2017, ch.3, pp. 207-322.

[7] R. J. LeVeque, *Finite difference methods for ordinary and partial differential equations*, SIAM, 2007

[8] D. Yan, F. P. Dawson, M. Pugh, and A. A. El-Deib, "Time-Dependent Finite-Volume Model of Thermoelectric Devices," IEEE Transactions on Industry Applications, vol. 50, no. 1, pp. 600-608, 2014.

Low-Cost ESP32, Raspberry Pi, Node-Red, and MQTT Protocol Based SCADA System

Atefeh Zare
Department of Electrical and Computer Engineering, Faculty of Engineering and Applied Science
Memorial University of Newfoundland (MUN)
St. John's, NL A1B 3X5, Canada
azare@mun.ca

M. Tariq Iqbal
Department of Electrical and Computer Engineering, Faculty of Engineering and Applied Science
Memorial University of Newfoundland (MUN)
St. John's, NL A1B 3X5, Canada
tariq@mun.ca

***Abstract*— Growing energy cost and demand has motivated many organizations to achieve smart ways to monitor, control, and save energy. Smart automation can reduce costs while still satisfying energy demand. The residential, commercial, and industrial sectors can utilize the technologies of the Internet of Things (IoT) to manage energy consumption better. This paper presents a low-cost, open-source, and reliable Supervisory Control and Data Acquisition (SCADA) system for home monitoring and control system. The presented SCADA system consists of analog sensors, ESP32, Node-RED, and Message Queuing Telemetry Transport (MQTT) through local Wi-Fi to remotely access and control appliances. This system helps the users to monitor various conditions in the home, such as temperature, humidity, pressure, and light intensity. Thus, users can remotely monitor various devices such as lights, fans, heating/cooling systems, make decisions based on the feedback of sensors.**

***Index Terms*—ESP32, Internet of Things (IoT), Wi-Fi network, MQTT, Node-RED, Sensors**

I. Introduction And Literature Review

Nowadays, the fraction of automation systems in the residential sector is rising rapidly because of numerous life improvements such as comfort, convenience, centralized control of appliances, cost reduction, energy-saving, security, and safety. As a home automation system enhances the life quality for users, especially for the elderly, differently abled persons, and people who want to monitor and manage their home devices' operation, it is quietly important to have a proper control system. SCADA technologies implement a unique platform that senses and monitors remote devices, acquires data from them, and sends limited control instructions. Besides that, this system allows users to discover the status of devices and their residence conditions remotely.

Home automation systems using SCADA consists of four major parts. The first part is the sensing devices placed at several locations throughout the home to measure the desired variables. The second part is Remote Terminal Units (RTUs) for acquiring remote data from sensors. The third one is Master Terminal Units (MTUs) to process the received data and deal with Human Machine Interfaces (HMI) [1]. The last part of SCADA is the communication channel to connect the RTUs to the MTUs [2 ,3].

SCADA technology has been developed over the past years to remotely monitor and control processes. In this work, an open-source and low-cost SCADA system based on the Internet of Things is introduced. This SCADA system utilizes reliable and commonly available components to fulfill the four main functions of a SCADA system: Data collection, network data communication, data presentation, and remote monitoring and supervisory control [4].

Several researchers around the world have designed SCADA systems based on IoT architecture in the past. Lekic et al. [5] performed an IoT-based SCADA system with the Raspberry Pi3. Temperature and humidity sensors acquired the desired data, and the IBM Bluemix cloud platform was utilized to receive, visualize, and manage the collected sensor data via the web while using the Node-Red and Web Socket protocols for data exchange and communication between the cloud platform and sensors connected to the Raspberry Pi.

In [6], Rai et al., Using IoT, implemented a system to provide the end-users with a cost-effective and portable intelligent monitoring system and enable devices remotely. The proposed system required a low-resolution video camera module interfaced with the Arducam ESP32 UNO board. The resulting video was streamed using ESP32 integrated Wi-Fi to display on the 1.8-inches SPI TFT Adafruit display module.

M. Al-Kuwari et al. [7] demonstrated sensing and monitoring platform on smart and IoT-based home automation system, where the authors presented a basic concept of how home automation can be deployed using IoT.

Pravalika et al. [8] implemented a home automation system through the Wi-Fi module, Massachusetts Institute of Technology (MIT) app, and a web page server using ESP32 to monitor home devices.

In [9], the MQTT-based home automation system using ESP8266 Node MCU was established for remote monitoring and controlling through a standard gateway. This system was designed utilizing the GSM (Global System for Mobile Communications) network.

In [10], Kodali demonstrated IoT-based smart environment monitoring using Arduino and C embedded programming devices, which were remotely controlled with the Internet. Temperature, humidity, light level, vibration, and air quality sensors connected to the controller for measurement and ESP8266 Node MCU for the Wi-Fi network.

Pujari et al. [11] demonstrated a smart home system that uses Wi-Fi to connect to the Internet, remotely monitor and control the appliances, and surveillance. Sensors were connected to ESP32 to collect data. All acquired data was

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

uploaded to Firebase via the ESP32's built-in Wi-Fi, making it possible to control the home environment using the developed applications.

In [12], the IoT-based SCADA system proposed which the Raspberry Pi3 and Intel Edison board used for acquiring sensor inputs. Collected sensor data are transmitted to Amazon Web Services (AWS) through MQTT brokers in Node-Red. On the AWS IoT platform, several monitoring and control schemes inducted using Amazon's Voice Service named Alexa.

In this work, a low-cost and open-source SCADA system is developed that uses Wi-Fi, the MQTT protocol, and Node-Red. The proposed system can provide detailed measurements of temperature, pressure, humidity, and light density. Besides that, the states of the devices at home are being traced. Therefore, users can understand the status of their home devices such as lights, fans, air conditioners, or heating/cooling systems then remotely control them.

II. System Design

The proposed SCADA system block diagram is presented in fig.1. It is implemented using ESP32 as the sensor gateway and RTU and Raspberry Pi2 as a local server. SparkFun Atmospheric Breakout Sensor-BME280 sensor and an LDR are connected to ESP32 to acquire the desired data and the dashboard to receive, visualize and manage the acquired sensors data over the web while using Node-Red and MQTT protocol for data exchange and communication between the MQTT clients such as end-users and the ESP32 connected sensors while the acquired data are being stored in SQLite as a web server. Each component is described in detail as following:

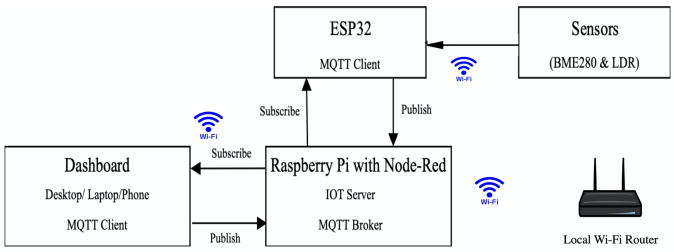


Fig. 1. The proposed SCADA system block diagram.

A. MQTT Communication Protocol

As shown in Fig. 2, MQTT is a publish/subscribe protocol. It is a lightweight and straightforward messaging tool designed for unreliable networks and constrained devices and low bandwidth. It is a suitable solution for this design since it provides simple communication between the server (MQTT broker) and clients (ESP32 microcontroller and computers and mobile devices) [9-10, 13]. Clients can subscribe to the topics or publish the data to topics of any kind of data. The broker then distributes the data to any client that has subscribed to that topic. Eclipse Mosquitto software is being run as a broker on the local server (Raspberry Pi2).

In this project, the ESP32 microcontroller publishes the sensors data with specific topics to the MQTT broker (Mosquitto broker), while personal computers and mobile devices subscribe to the topics to visualize the published data on the server. MQTT protocol makes the supervisory control possible in this project.

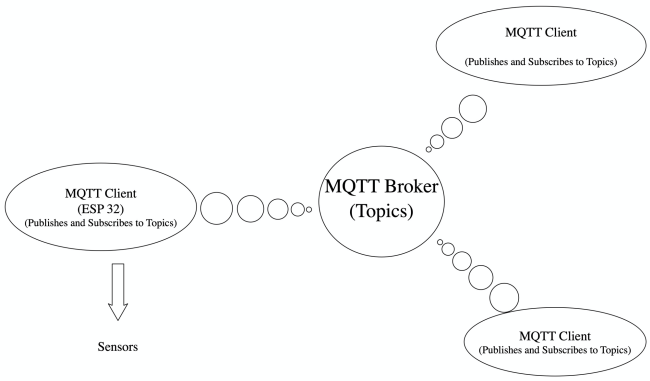


Fig. 2. MQTT Architecture.

B. Raspberry Pi

The Raspberry Pi2 model B is used as the local server at a local network, including MQTT broker, Node-Red, SQLite.

C. ESP32 Thing

The ESP32 Thing is developed by SparkFun Electronics. It is one of the most low-cost, low-power, and small microcontrollers included Wi-Fi and Bluetooth modules. Fig. 3 shows a picture of the SparkFun ESP32 board.

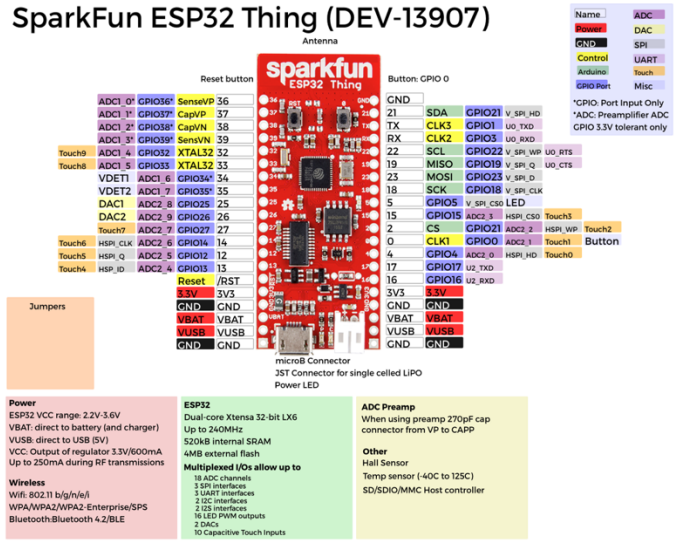


Fig.3. Pin description of ESP32.

The board operating signal voltage range is 2.2 V to 3.6 V, and it can be supplied with either a 5 V USB power supply or with a battery. The microcontroller is programmed with the Arduino software integrated development environment (IDE). The Arduino IDE allows experts to write the desired programs and upload them to the board via a USB cable. In this work, the ESP32 microcontroller is programmed as an MQTT client using the Arduino IDE software. The MQTT

Client Library called PubSubClient. The program is such that the board collects the measured real-time values of desired variables, displays the values on the Arduino IDE Serial Monitor, and continuously publishes real-time data to the MQTT broker.

D. SparkFun Atmospheric Sensor Breakout BME280

The SparkFun BME280 Atmospheric Sensor Breakout is used to measure pressure, humidity, and temperature readings, all with a tiny breakout. It is shown in Fig. 4. The BME280 Breakout is designed to be deployed in indoor/outdoor navigation, home automation, etc. The on-board BME280 sensor measures atmospheric pressure, humidity, and temperature. The breakout presents a 5V tolerant I2C interface (with a pull-up resistor to 3.3V to be cooperative with ESP32).

Fig.4. SparkFun Atmospheric Sensor Breakout BME280.

E. Light Dependent Resistor (LDR)

The LDR is a variable resistor. Fig. 5 presents LDR and its step-down resistor connection. The resistance of the LDR inversely proportionates to the light intensity. It shows maximum resistance in the absence of light and minimum resistance in the presence of light. Considering this characteristic, in this project, during the nighttime, LDR turns the LED ON and reversely makes the LED OFF during the daytime.

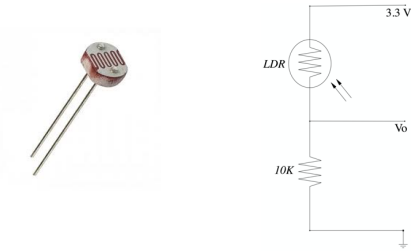


Fig.5. LDR and its step-down resistor connection.

The relationship between the resistance RL and light intensity Lux for a typical LDR is [14]:

$$R_L = 500/Lux \quad (1)$$

With the LDR connected to 3.3V through a 10K resistor, the output voltage of the LDR is:

$$V_O = (10*R_L)/(R_L+10) \quad (2)$$

From equation (1) and (2), light intensity obtained as follows:

$$Lux = (500/V_O)-50 \quad (3)$$

The obtained Lux value is being used for controlling the LED.

F. Node-Red

Node-RED is an open-source programming tool used for wiring together hardware devices, APIs, and online services smartly [5]. It can be installed on a Linux-based platform, and it provides a browser-based editor that makes it easy to wire together flows using various nodes in the palette that can be deployed to its runtime.

G. Database and dashboard

An SQLite database using the litedb node on the Node-RED platform is installed to create database tables and store data. As a very lightweight relational database, SQLite does not need complex setup procedures, making it an ideal database management system to use for embedded systems and rapid prototyping. In this work, SQLite is set up to generate database, then stored data an easy to use dashboard (Node-Red dashboard) is used to read the acquired data from the data base.

Since the sensor data are constructed data type as MQTT messages that will be published to an MQTT broker, Arduino IDE sketch and Node-Red flow are used to transfer those MQTT messages and then stores them in the database. Node-Red dashboard is a web-based data monitoring tool that can be combined to SQLite.

III. System Testing And Results

The home automation system's proposed design is used to monitor and control the house appliances while maintaining the minimum expected comfortable living conditions.

The flow chart of the proposed system is shown in Fig. 6. Accordingly, a prototype is shown in Fig. 5 is made to measure the desired variables. This prototype can be located in each room of the house. The sensors acquired data is being processed and used to control smart appliances such as smart LEDs, air conditioners, fans, etc.

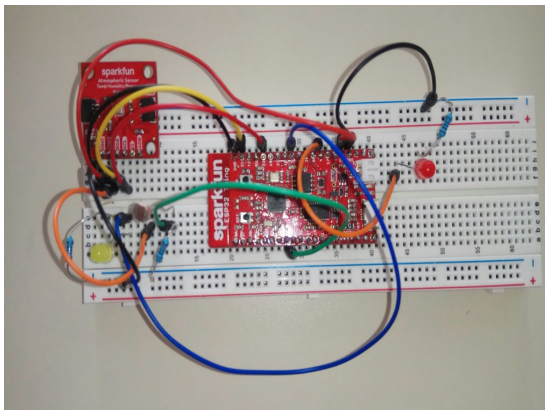


Fig. 5. Proposed SCADA system prototype.

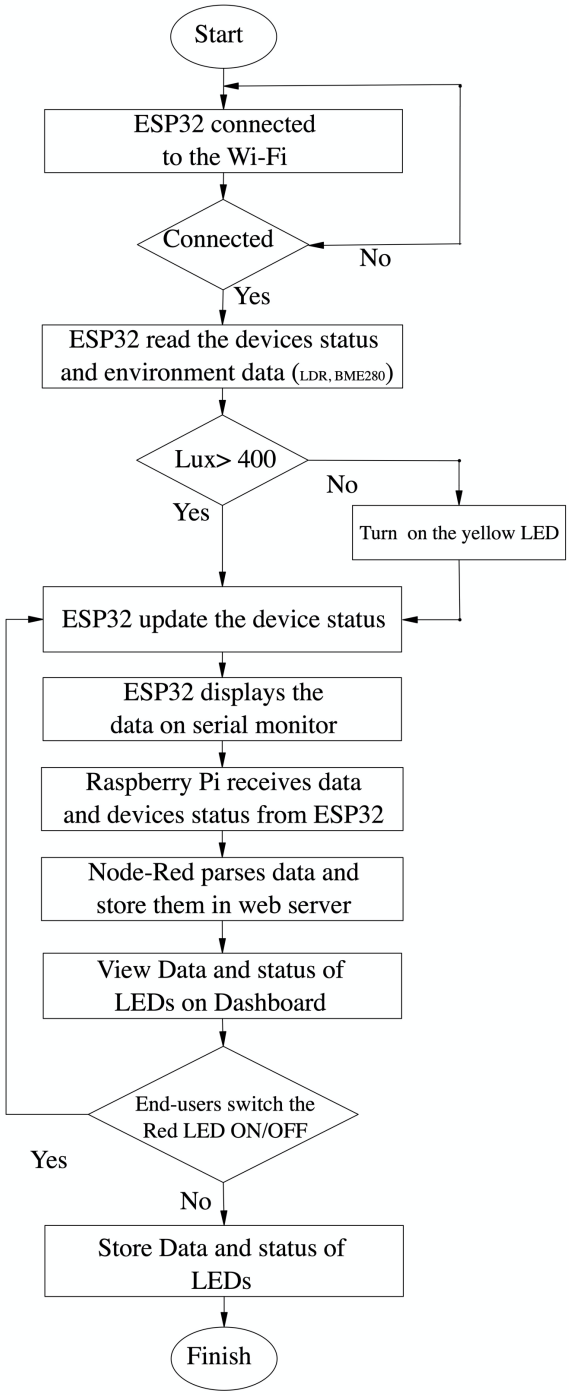


Fig. 6. The flow chart of the proposed SCADA system.

In this project a yellow LED is used to test the automatic control ability of the system and a Red LED is implemented to evaluate the supervisory control capability of the proposed SCADA system.

This system monitors light intensity, temperature, humidity, and pressure values. If the recorded light intensity value goes beyond the set high or low value of the threshold level, the controller will generate the signal to turn the LED OFF/ON. Then the user will be notified by the LED status on the dashboard, this feature is presented in Fig. 7 and 8.

As shown in Fig. 7, the LDR is darken using a piece of paper to change the state of the LED which is presented in Fig. 8.

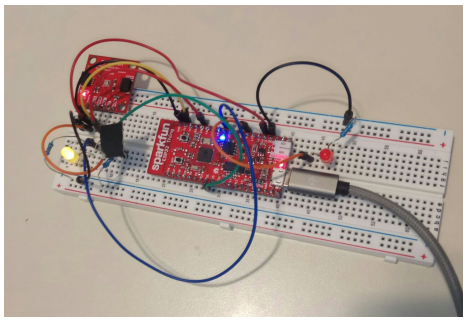


Fig. 7. The LED is turned on automatically based on the threshold value.

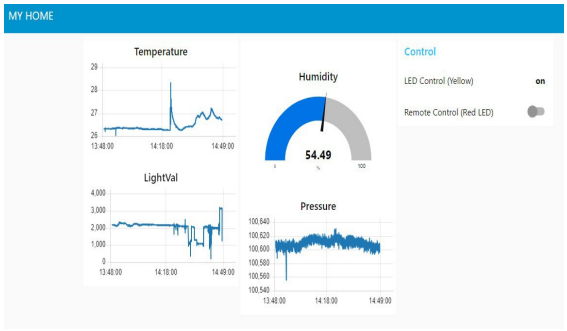


Fig. 8. The status of LED in Node-Red dashboard.

As mentioned earlier, the ESP32 controller subscribes to the published command, created by the end-user through the web interface, and transmitted by the Node-Red. Based on the received message, ESP32 sends a command to the LED to turn it ON/OFF

The implemented system is tested to evaluate its capability. Results provided in Fig. 9, 10 show the recorded values of temperature, humidity, and light density, and pressure, which are acquired and displayed over time through a data-viewing dashboard.

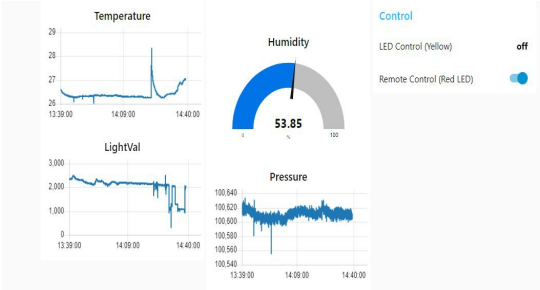


Fig. 9. The Red LED is controlled remotely by using Node-Red dashboard.

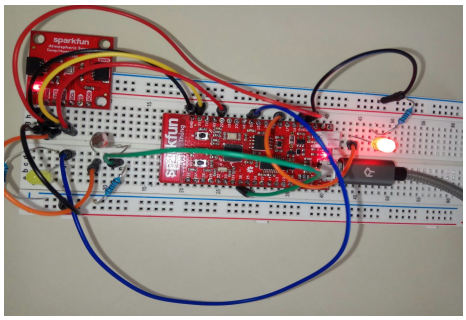


Fig. 10. The user remotely turned the Red LED on.

IV. Conclusion

The design of a low-cost and flexible and reliable IoT home automation system is presented. This proposed SCADA system is suitable for monitoring and controlling multi devices at home automatically and remotely. The system works in three phases. In the first one, the system monitors the temperature, pressure, humidity, and light intensity and uploads the data to database storage (SQLite) by using MQTT protocol and Node-Red installed on Raspberry Pi, then to the database connected dashboard. In the next phase, the system automatically controls the LED when the light intensity level exceeds the high or low predefined value of light intensity using a microcontroller (ESP32). In the last one, the supervisory control is achieved; the end-user can monitor and control the home devices through the web interface or mobile application. Once the message is sent through Node-Red along with the MQTT protocol to the microcontroller, it will be executed by ESP32 and turns the devices ON/OFF. The developed system cost is low, simple to operate, and is simply embedded with home devices.

Acknowledgment

The authors would like to thank the School of Graduate Studies, Faculty of Engineering and Applied Science, Memorial University and Pooya Pajouhan Nirouye Golpa Company.

References

[1] Lamine and H. Abid, "Remote control of domestic equipment from an Android application based on Raspberry Pi card", IEEE transaction 15th international conference on Sciences and Techniques of Automatic control & computer engineering - STA'2014, Hammamet, Tunisia, December 21-23, 2014. [Accessed 21 September 2014].

[2] B. S.S. Tejesh and S. Neeraja, "A Smart Home Automation system using IoT and Open Source Hardware", International Journal of Engineering & Technology, vol. 7, no. 27, p. 428, 2018. Available: 10.14419/ijet.v7i2.7.10856.

[3] L. Aghenta and M. Iqbal, "Low-Cost, Open Source IoT-Based SCADA System Design Using Thinger.IO and ESP32 Thing", Electronics, vol. 8, no. 8, p. 822, 2019. Available: 10.3390/electronics8080822.

[4] A. Sajid, H. Abbas and K. Saleem, "Cloud-Assisted IoT-Based SCADA Systems Security: A Review of the State of the Art and Future Challenges", IEEE Access, vol. 4, pp. 1375-1384, 2016. Available: 10.1109/access.2016.2549047.

[5] M. Lekić and G. Gardašević, "IoT sensor integration to Node-RED platform," 2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH), East Sarajevo, 2018, pp. 1-5, doi: 10.1109/INFOTEH.2018.8345544.

[6] P. Rai and M. Rehman, "ESP32 Based Smart Surveillance System," *2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, Sukkur, Pakistan, 2019, pp. 1-3, doi: 10.1109/ICOMET.2019.8673463.

[7] M. Al-Kuwari, A. Ramadan, Y. Ismael, L. Al-Sughair, A. Gastli and M. Benammar, "Smart-home automation using IoT-based sensing and monitoring platform," *2018 IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG 2018)*, Doha, 2018, pp. 1-6, doi: 10.1109/CPE.2018.8372548.

[8] V. Pravalika and Ch. Rajendra Prasad, "Internet of Things Based Home Monitoring and Device Control Using Esp32", International Journal of Recent Technology and Engineering (IJRTE), vol. 8, no. 4, pp. 58-62, 2019.

[9] R. K. Kodali and S. Soratkal, "MQTT based home automation system using ESP8266," *2016 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, Agra, 2016, pp. 1-5, doi: 10.1109/R10-HTC.2016.7906845.

[10] R. K. Kodali and S. Soratkal, "MQTT based home automation system using ESP8266," *2016 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, Agra, 2016, pp. 1-5, doi: 10.1109/R10-HTC.2016.7906845.

[11] U. Pujaria, P. Patil, N. Bahadure and M. Asnodkar, "Internet of Things based Integrated Smart Home Automation System", SSRN Electronic Journal, 2020. Available: 10.2139/ssrn.3645458.

[12] A. Rajalakshmi and H. Shahnasser, "Internet of Things using Node-Red and alexa," *2017 17th International Symposium on Communications and Information Technologies (ISCIT)*, Cairns, QLD, 2017, pp. 1-4, doi: 10.1109/ISCIT.2017.8261194.

[13] K. Chooruang and P. Mangkalakeeree, "Wireless Heart Rate Monitoring System Using MQTT", Procedia Computer Science, vol. 86, pp. 160-163, 2016. Available: 10.1016/j.procs.2016.05.045.

[14] D. NAGARAJU, C. KIREET, N. KUMAR and R. JATOTH, "Performance Comparision Of Signal Conditioning Circuits For Light Intensity Measurement", World Academics Journal of Engineering Sciences, vol. 01, no. 02, pp. 2007 (1-10), 2014. Available: 10.15449/wjes.2014.2007.

Clock Tree Optimization of FPGA Semi-Custom Memory with SEU FlipFlops

Sourabh Aditya Swarnkar
Memory Design Group
Xilinx Pvt. Ltd
Hyderabad, India
sourabh@xilinx.com

Kumar Rahul
Memory Design Group
Xilinx Pvt. Ltd.
Hyderabad, India
kumarr@xilinx.com

Mohammad Anees
Memory Design Group
Xilinx Pvt. Ltd.
Hyderabad, India
manees@xilinx.com

Santosh Yachareni
Memory Design Group
Xilinx Pvt. Ltd
Hyderabad, India
santoshy@xilinx.com

Abstract— Flip-flops (FFs) and memory (including Block RAM and Configuration RAM) are the key elements in Field Programmable Gate Arrays (FPGA). A single radiation event can flip the storage node of the sequential elements. FPGAs are widely used in radiation environments such as space, the mitigation of single event upset (SEU) in SRAM based FPGAs is increasingly important. Generally, the SRAM macros implement well optimized in-built Error Correction Codes (ECC). For SEU tolerance, traditional FFs are replaced by SEU hardened FFs. SEU FFs have higher setup and clock-to-out delays and degrades the performance of SRAMs in FPGAs.

High performance SRAMs in FPGAs are implemented using Hybrid approach of custom and PNR (Place and Route). FPGA applications demanding more features such as ability to form deeper memories, and the programmable word size etc. This leads to increase in the logic and wirelengths, which further introduces higher delays in designs and impacts performance.

So, the optimal clock-distribution network in high density memories along with SEU tolerance features, is one of the key aspects of high-speed SoC designs. This paper demonstrates methods to implement high quality clock tree to mitigate delay penalties introduced by the SEU hardened FFs and higher wirelengths. Experimental results demonstrate that the proposed approach significantly improves clock tree performance.

Keywords— SEU (Single Event Upset), PNR (Place and Route), ECC (Error Correction Code), SRAM (Static Random-Access Memory), CTS (Clock Tree Synthesis), FF (Flipflop), FPGA (Field Programmable Gate Array), ECC (Error Correction Code)

I. Introduction

In deep submicron technologies, circuit node capacitance and supply voltage decrease considerably. SEU impact is quite high on nodes with very small capacitances, and small charge storing capabilities. SEU [1]- [5] might be triggered by multiple reasons but the main source is cosmic rays, or Ţ-particles originated by the decay of uranium and thorium impurities presented within packages [6]-[8]. These charges get collected by drain/source diffusions, which can alter the logical value of storage devices.

The SEU sensitivity of SRAMs in FPGA depends on in-built ECC [9] [10], RAM bitcells, and FFs. In-built ECC [9] [10] improves the SEU tolerance of the SRAMs in FPGA. The circuit reliability is greatly improved by replacing the traditional FFs by SEU hardened FFs. SEU hardened FF designs [11] [12], use techniques of strengthening latch designs to increase SEU tolerant capabilities. SEU hardened FF designs [11] [12] usually employ additional locking loops to restore the original state for a TF (Transient Fault) impacted nodes (Fig. 1). With the additional loop, along with the existing Q & Qb nodes, SEU FF has an extra pair of QQ & QQb nodes as shown in Fig.1. These nodes are placed in layout in this sequence: Q, Qb, QQ, QQb. The concept is to replicate the Q and Qb nodes spaced apart in the layout. Even if one of the nodes gets corrupted due to SEU, then the other three nodes restore the corrupted nodes. SEU FFs have a larger area compared to the traditional FFs and the capacitance of internal nodes is high. Hence clock-to-out, setup delays of the SEU FFs are higher. The design goal is to mitigate these higher clock-to-out and setup delays by reducing the clock insertion delays of clock trees, without degrading the clock skews.

Delays are highly impacted by the big wirelengths in deep submicron technologies. The delay due to the wirelength (RC impact) of the clock tree network can be reduced by placing the FFs close to each other. Previous work has been proposed to reduce the RC delay register placements [14]. These placement methods are applied before the clock tree synthesis (CTS), and the CTS tool may not be aware of the clusters. This will cause registers inside one cluster to be driven by different buffers. One significant work to handle this issue is from Jianchao Lu et al. [15]. It involves incremental register placement into the CTS stage so that the CTS tool is aware of the clusters. The skew requirement of the clock tree can be strictly maintained as the limited register placement. This solution is easy to plug into the mainstream PNR flow. So, the benefits obtained from industrial physical synthesis tools are preserved and clock tree timing is still guaranteed.

This paper implements the above solution to achieve the careful clustering of the SEU hardened FFs. We further introduce a new approach to visualize the results using python 3D plots. With this approach, we can quickly analyze the results for FFs placement and corresponding clock sink delays. It makes a comparison of results among the various experiments very effective and more conclusive.

II. Design Challenges

A. Data & clock paths delays

System-level frequency requirements for complex memories along with multiple functional modes are tricky to meet. Each functional mode has various flavors of data path delays. Access time is the time needed to access the data from the memory and perform some operation on the data. Functions possessing the bigger delays in timing paths, are of the primary concern to meet the overall system performance.

With the introduction of hardened flops, system-level clock-to-out, setup times, and access times are degraded. SEU hardened flops, with extra RC delays in internal nodes, have higher clock-to-output delays. Various experiments

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

[13] show that hardened latches have 40% extra delay penalty compared to un-hardened latches.

To reduce the access time of memories, faster clock-to-out is the key performance factor. Both data and clock path delays need to be minimized without impacting setup and hold requirements. Hence, clock tree implementation with accurate placement of hardened flops to achieve optimum clock routing delays is very critical in terms of keeping minimum and identical insertion delays across various FPGA functional modes inside memories.

B. *Prioritizing the critical functional modes for timing closure*

Analysis of the different functional modules and their interaction plays an important role in budgeting the clock tree timings. Larger skews between various flops of different functional blocks can result in delay penalty. So, all the functions in memory need to be assessed to evaluate any dependencies on each other as shown in Fig. 2. This analysis is very important to restrict the registers inside one cluster to be driven by different buffers.

For example, memories to have faster access times, read paths need to be further optimized. Having features like cascading option to connect multiple memories to form a deeper single memory, will increase the overall access time of the memory. To handle this, various pipeline stages are added between these features and the SRAM block, and to reduce the combinational logic delay between them. Reduction in combinational logic delay will reduce the total time-period required, hence increasing the operating frequency. Pipelining involves the insertion of FFs between combinational logic blocks, this reduces glitch propagation effectively [17]. These pipelining flops are also implemented with SEU hardened FFs.

Similarly, write path with data input being not in critical path, enables to have more delay without impacting performance. But functions, which fall under the critical timing paths such as read paths, impacting system frequency, are desired to have optimum clock delays. Memory access time is the dominating performance factor. So, the read path being critical for the timing, we can further partition the read paths based on various functions in the design. For simplicity, we will refer to this partitioning as different modules like 'Read_path1', 'Read_path2', 'write path' etc. in the upcoming sections.

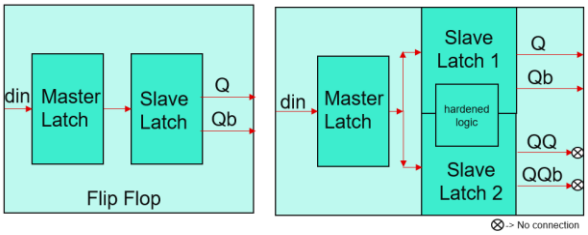


Fig. 1 (a) Standard FF (b) Hardened SEU FF

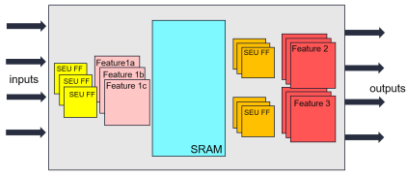


Fig. 2 Conceptual memory with programmable features, SEU hardened flops and SRAM array macro

III. Review Of Standard CTS Methods

Manual synthesis of a synchronous clock tree [16] is a very complex and labor-intensive process, which is obviously not applicable in modern designs if we take into consideration the volume of modern projects and the complexity.

For automatic CTS [16], the clock tree buffer is built according to the frequency specifications and slew requirements. CTS generates the clock tree topology, and balance the clock phase delay with appropriately sized inserted clock buffers. Fig. 3 represents the automatic CTS mode.

Regardless of what kind of methodology is used for the synthesis of a CTS, the main goal is to minimize the access time of memories, and hence faster clock-to-out is the key performance factor. Clock-to-out has two main components: ***Clock insertion delay*** **(D_{cin})**, ***hardened SEU flop delay*** **(D_{SEUFF})**. Hardened flop delay cannot be optimized by CTS. So, during clock tree implementation, the additional 40% delay penalty of hardened flops need to be compensated. With the accurate placement of the hardened flops, both data and clock path delays will be optimized without impacting setup and hold requirements.

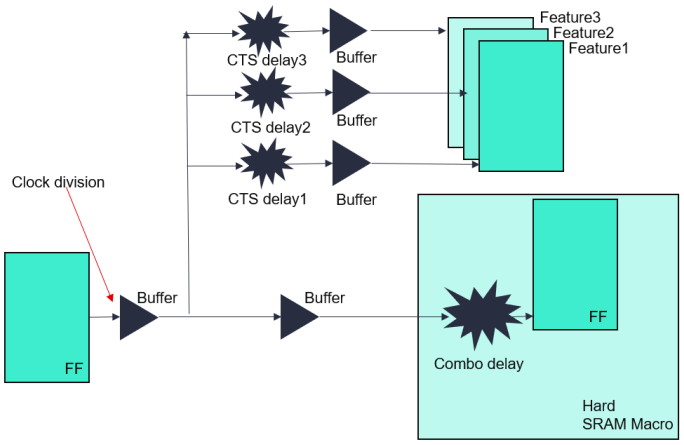


Fig. 3 Representations of Automatic CTS Mode

IV. Proposed Custom Approach

The memory Macro and periphery are implemented in 7nm FinFET technology, using an automatic PNR tool. The block consists of about 15000 gates, including 750 hardened SEU FFs and a couple of hard Macros. Before our proposed method is introduced, let's consider a series of CTS experiments for 'Read_path1' as targeted module discussed in section II/C.

Method 1

In the initial experiment, we synthesize the clock tree in automatic CTS mode discussed in section III. The slowest clock-to-out delay within the targeted module was 274ps and the overall slowest clock to out was 302ps. The maximum skew reported was 85ps.

Method 2

This experiment includes clock balancing using early delay option available in automatic CTS mode. This option applies the exception for early/shortest path calculations for the targeted clock sinks. The targeted module was

constrained to have an early delay of 80ps. With this approach, the slowest clock-to-out in the targeted module improved (w.r.t. Method 1) to 218ps. But it degrades the overall worst clock-to-out delay (w.r.t. Method 1) to 330ps. Also, the skew degraded significantly to 130ps.

Method 3

In this experiment, we used the Clock Grouping technology available in automatic CTS mode as shown in Fig. 4. The targeted module's clock sinks were kept in a separate clock balancing group. With this approach, the slowest clock-to-out delay improved to 245ps (w.r.t. Method 1). There is no significant improvement in the overall worst clock-to-out delay and clock skew (w.r.t. Method 1).

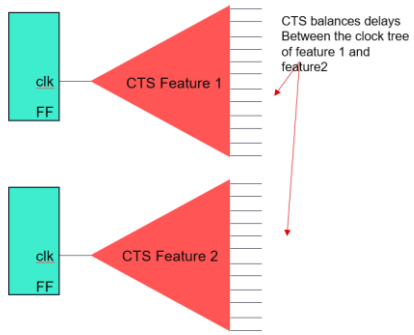


Fig. 4 Graphical Representation of clock grouping in Auto CTS

Method 4

This experiment is the extension of Method 3, where we not only used the Clock Grouping but also use the placement bounds to cluster all the hardened SEU FFs of one targeted module to be physically placed together.

Using placement bound constraints, we physically clustered the placement of all hardened flops of the targeted feature. This ensures minimizing RC delays of clock and data routing. The cells placed in these bounds are not disrupted for placement by other features/logics.

With this approach, the worst clock-to-out delay significantly improved and reduced to 200ps, while overall worst clock delay also reduced to 250ps. The maximum skew is also reduced to 70ps. (w.r.t. Method 3).

Method 5

This final experiment is further fine-tuning of Method 4, with the following additional options:

A. Concurrently generate clock network and incrementally perform register clustering to reduce the clock tree wirelength, i.e., incremental placement legalization of clustered flops.

B. Perform register clustering to reduce the clock tree wirelength.

C. Targeted both read path modules ('Read_path1' & 'Read_path2') and write path features to be clustered in three separate placement bounds. The clock sink pins for these modules are kept in separate clock grouping as shown in Fig. 5.

D. Three separate placement bounds were defined:

Placement bound 1 for 'Read_path1' included 150 SEU FFs. Placement bound 2 for 'Read_path2' included 150 SEU FFs. Placement bound 3 for 'write path' included 148 SEU FFs.

E. Both 'Read_path1' & 'Read_Path2' were given the flexibility to choose the buffers from faster libraries as well. Hence adding the hybrid flavors of CTS buffers.

F. All other clock sinks were left for automatic CTS to place and optimize automatically.

This approach significantly improves the worst clock-to-out delay to 174ps for both targeted Read path Features. The overall worst-case clock delay improved to 209ps and the maximum skew also further reduced at 54ps.

Overall comparison of the results across the various methods is captured in TABLE I.

TABLE I. STATISTICS OF CTS RESULTS ACROSS EXPERIMENTS

Method	*Worst D_{cin} + D_{SEUFF} Targeted Feature (ps)*	*Worst D_{cin} + D_{SEUFF} Overall (ps)*	*Maximum Skew (ps)*
1	274	302	85
2	218	330	130
3	245	290	79
4	200	250	70
5	174	209	54

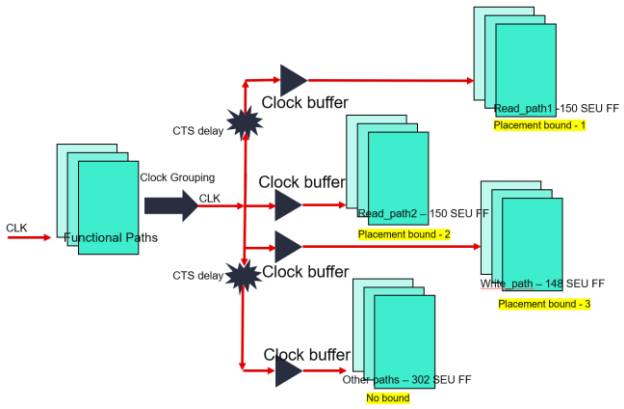


Fig. 5 Graphical Representation of clock grouping with placement bounds across the different features.

V. RESULT VISUALIZATION

Data visualization provides us with a quick, clear understanding of the results. We can visualize large volumes of data understandably and coherently, which in turn helps us comprehend the information and draw conclusions and insights. Standard PNR tools dump multiple report files. With the iterative approach, every experiment has a lot of samples of results, comparison of these results becomes very tedious.

We further introduce a new approach to visualize the results using python 3D plots. Automation of results to be processed into python script was used to get samples with every iteration. With this approach, analysis of the results for SEU FFs placement and corresponding clock sink delays were very quick. It made a comparison of results among the various experiments very effective and more conclusive.

Fig. 6 & Fig. 7 represents such 3D plots of physical FF locations in the die along with clock insertion delay (D_{cin}),

hardened SEU flop delay (D_{SEUFF}) associated with each FFs. X & Y represent the physical locations, while Z-axis represents the clock insertion delay D_{cin}+ D_{SEUFF}.

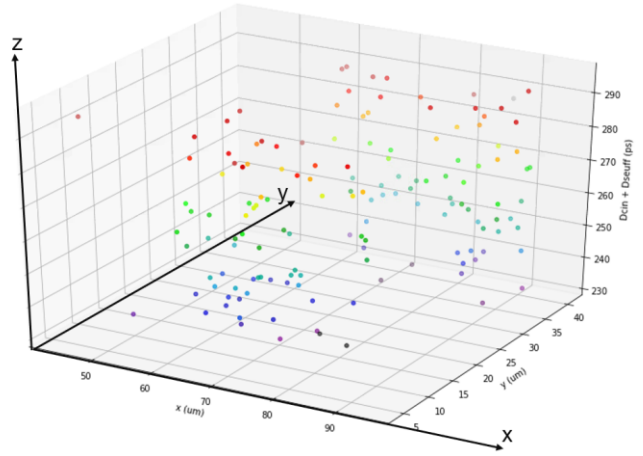


Fig. 6 'Read_path1' all SEU FF physical location Vs Clock insertion Delay representation before proposed implementation (Method 1)

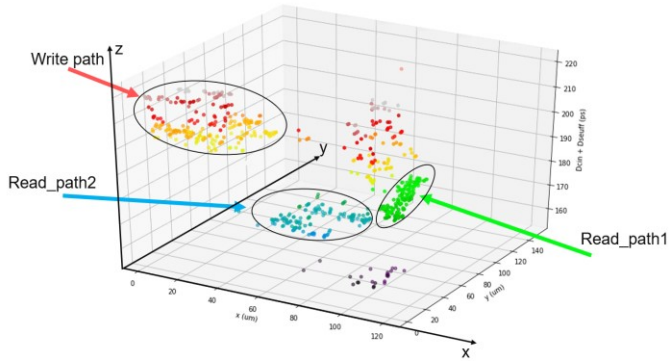


Fig. 7 All FF physical location Vs Clock insertion Delay representation after proposed implementation (Method 5)

From Fig. 6, we can visualize the Method 1 results, all the 'Read path1' related SEU FFs were not only scattered along with the die but also had a higher insertion delay. We can also confirm from the plots that more than 50% of the SEU FFs had (D_{cin}+ D_{SEUFF}) number higher than 240ps.

From Fig. 7, we can visualize the Method 5 results, which include the related FFs of all the features. We can observe that with our approach, SEU FFs from targeted features are clustered together. This kind of FFs placement ensures to minimize local skew within a specific functional module. From the plot, we can conclude that both wirelengths along with insertion delays to sink pins are reduced. This also helped improve (D_{cin}+ D_{SEUFF}) delay of all the other clock paths as well.

VI. Conclusions

The in-built ECC [18] and the SEU FFs considerably increase the reliability of the SRAMs in FPGAs. The proposed method performs simultaneous CTS and incremental register placement to minimize the total clock wirelengths. The method has the advantage of not significantly changing the existing placement, which has been optimized for multiple factors, such as network congestion and manufacturability. Designing high-density custom memories with multiple functions have longer turnaround design times. The proposed experiments significantly reduce the effort needed for timing fixes, hence faster turnaround times. This solution is easy to plug into the standard automated PNR flows.

References

[1] Eishi Ibe, Hitoshi Taniguchi, Yasuo Yahagi, Ken-ichi Shimbo, Tadanobu Toba, "Impact of Scaling on Neutron-Induced Soft Error in SRAMs from a 250 nm to a 22 nm Design Rule", IEEE Transactions on Electron Devices (Volume: 57, Issue:7, July 2010)

[2] A. Klockmann, G. Georgakos, M. Goessel, A new 3-bit burst-error correcting code", On-Line Testing and Robust System Design (IOLTS), 2017 IEEE 23rd International Symposium

[3] Sanghyeon Baeg, Ansan, ShiJie Wen, Richard Wong, "SRAM Interleaving Distance Selection With a Soft Error Failure Model", IEEE Transactions on Nuclear Science (Volume: 56, Issue: 4, Aug. 2009)

[4] Anand Dixit, Alan Wood, "The impact of new technology on soft error rates", Reliability Physics Symposium (IRPS), 2011 IEEE International

[5] Jorge Tonfat, "Analyzing the Influence of the Angles of Incidence and Rotation on MBU Events Induced by Low LET Heavy Ions in a 28-nm SRAM-Based FPGA", IEEE Transactions on Nuclear Science (Volume: 64, Issue:8, Aug.2017)

[6] G. Messenger, "Collection of charge on junction nodes from ion tracks," IEEE Trans. Nucl. Sci, NS-29(6):2024–2023, December 1982.

[7] S. Kang and D. Chu, "CMOS circuit design for the prevention of single event upset," Proc. Of IEEE Int. Conf. On Computer Design, pp: 385–388, October 1986.

[8] C. M. Hsieh, P. C. Murley, and R. R. O'Brien, "Collection of charge alpha-particle tracks in silicon devices," IEEE Trans. On Electron Devices, ED-30:686–693, 1983.

[9] Kumar Rahul, Santosh Yachareni, "Area and power efficient ECC for multiple adjacent bit errors in SRAMs, " 2020 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, USA, 10.1109/ICCE46568.2020.9042979

[10] Kumar Rahul, Santosh Yachareni, "Low Power and Area efficient Forward Error correction code for SRAM in FPGAs, " 2020 8th International Electrical Engineering Congress (iEECON), Chiang Mai, Thailand, Thailand, 10.1109/iEECON48109.2020.229511

[11] Riadul Islam, "A highly reliable SEU hardened latch and high performance SEU hardened flip-flop, " Thirteenth International Symposium on Quality Electronic Design, Santa Clara, CA, USA, , 2012, 10.1109/ISQED.2012.6187516

[12] Ameet Chavan, Eric MacDonald, Joseph Neff, Eric Bozeman "Radiation hardened Flip-Flop design for super and sub threshold voltage operation, "2011 Aerospace Conference, Big Sky, MT, USA, 10.1109/AERO.2011.5747457

[13] M. Glorieux, S. Clerc, G. Gasiot, J. Autran and P. Roche, "New D-Flip-Flop Design in 65 nm CMOS for Improved SEU and Low Power Overhead at System Level," in IEEE Transactions on Nuclear Science, vol. 60, no. 6, pp. 4381-4386, Dec. 2013, doi: 10.1109/TNS.2013.2284604.

[14] Y. Cheon, P.-H. Ho, A. Kahng, S. Reda, and Q. Wang, "Power-aware placement," in Proceedings of the 42ndannual Design Automation Conference, 2005, pp. 795–800.

[15] Jianchao Lu and B. Taskin, "Incremental register placement for low power CTS," 2009 International SoC Design Conference (ISOCC), Busan, 2009, pp. 232-236, doi: 10.1109/SOCDC.2009.5423805.

[16] N. Avdalyan and K. Petrosyan, "Processing An Effective Method For Clock Tree Synthesis," 2019 IEEE East-West Design & Test Symposium (EWDTS), Batumi, Georgia, 2019, pp. 1-5, doi: 10.1109/EWDTS.2019.8884370.

[17] Hyeonmin Lim, Kyungsoo Lee, Youngjin Cho and Naehyuck Chang, "Flip-flop insertion with shifted-phase clocks for FPGA power reduction," ICCAD-2005. IEEE/ACM International Conference on Computer-Aided Design, 2005., San Jose, CA, USA, 2005, pp. 335-342, doi: 10.1109/ICCAD.2005.1560090.

[18] Kumar Rahul, Santosh Yachareni, "Encoding and decoding of information using a block code matrix," U.S. Patent 8 972 833, Mar. 3, 2015

Design of Universal logic gates and Majority Gate Using One clock pulse based CMOS Capacitor Coupled Threshold Logic

Arpita Dey, Mili Sarkar, Riddhi Roy, Tamonash Kanti Santra, G.S.Taki
Department of Electronics and Communication Engineering
Institute of Engineering & Management, Kolkata, India
email: mili.sarkar@iemcal.com, arpitadey060@gmail.com,
tamonashsantra@gmail.com,riddhiroyiem@gmail.com,gstaki@iemcal.com

Abstract **- In this era of rapid advancement in IC technology, multi-valued logic is playing a bigger role in incorporating multiple functions within single block. Threshold logic gate using Capacitor coupling logic is one of the most effective methods which has given fruitful results. This research paper presents one capacitor coupling logic (C3L) circuit with one clock pulse to implement the CMOS based threshold logic gates .All the simulations are carried out in TANNER software(T spice) using 250nm technology. Here NOR,NAND,AND,OR and majority gate have been designed using this CCLG(Capacitor Coupling Logic Gate) by changing the value of a reference voltage.**

Keywords**:** ***Threshold Logic, Capacitor coupling logic (CCL), Majority gate***

1. INTRODUCTION

Threshold Logic gate normally compares the sum of weighted inputs with a reference voltage to provide digital output [1]. It had been designed in many static and dynamic design style [2,3]. One of these dynamic design of TLG is Capacitor coupling logic.

In capacitor coupling threshold logic (CCTL) family all the input signals are coupled to one common transistor gate of a differential transistor pair through coupling capacitors. A reference signal is coupled to the other transistor gate of the transistor pair. The voltage difference between these two gates is sensed by a sense amplifier to generate two complementary logic outputs [5]. For CCTL gates with a large number of inputs, the transistor count, interconnections and parasitic capacitances along the signal path are significantly reduced, which results in high switching speed and small chip area. Furthermore, the DC power dissipation in CCTL gates is almost nil. We have modified the existing circuit by giving only one clock pulse and also input capacitors are removed for reference voltage to minimize the circuit complexity.

A majority gate is a logical gate which is used in many complex Boolean circuits. A majority gate gives 1 if and only if more than 50% of its inputs are true. We have designed majority gate using proposed CCTL which is further used for Full Adder [6, 7] circuit. As the basic CCTL circuit is providing C_{out} and its complement (C_{out}') which is further used in Sum Circuit.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

2. CAPACITOR COUPLING LOGIC

In capacitor coupling logic [7,8,9,10,11] both P-MOS and N-MOS are used. They are used in two ways-as CMOS switching circuit and as differential pair circuit. A single clock pulse is used which synchronizes the entire circuit. Two outputs are available from the circuit (Out1 & Out2) which are in complemented form. The basic circuit is shown in Figure1.

When the circuit is in initial stage i.e clk=0 both the outputs become V_{dd} momentarily. The Gate-source voltage of N-MOS (N3 &N4) becomes high. As a result, both the output becomes 1. Then the differential circuit comes into play. The differential circuit compares the Gate voltages of N1 & N2 N-MOS. At N1 the input binary combinations are applied through capacitors and at the gate of N2 reference voltage is directly applied. If $V_{gate\ N1} > V_{gateN2}$ then N1 will be on and N2 will be off. If the reverse scenario happens then the case will be vice versa.

Now if $V_{gateN1} - V_{gateN2} = +\Delta V > 0$ then Out1 is evaluated to low and Out2 becomes high. If $V_{gateN1} - V_{gateN2} = -\Delta V < 0$ then Out1 becomes high and Out2 becomes low. ΔV can be evaluated very easily with the help of voltage divider rule and superposition theorem.

$$\Delta V_{gateN1} = \frac{C_A}{C_A + C_P} V_A + \frac{C_B}{C_B + C_P} V_B \quad \text{------- [1]}$$

$$\Delta V_{gateN2} = V_{ref} \quad \text{---------- ------- [2]}$$

$$\Delta V = \Delta V_{gateN1} - \Delta V_{gateN2} \text{-------- -------[3]}$$

V_{gateN1} and ΔV_{gateN2} is given in eqn 1 and 2. Here, Cp is the total parasitic capacitance of N1 N-MOS gate junction node.Parasitic capacitance is an unavoidable and unwanted capacitance which basically exists between the metal contacts and source or drain junction. During fabrication process due to high temperature the metal contacts in gate region expand and reaches at the vicinity of source and drain region. As a result, a non-negligible parasitic capacitance arises. As more than one input has been provided in the N1 gate terminal so a considerable amount of parasitic capacitance comes into play. For other terminals it is not so significant as only one voltage input terminal is present there.

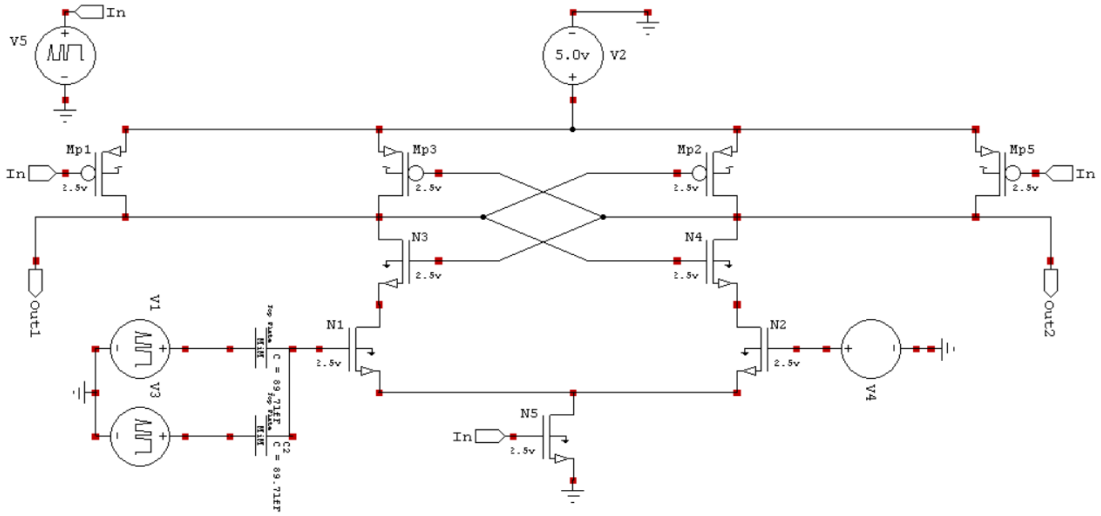


Fig.1. Modified CCCL based NAND gate

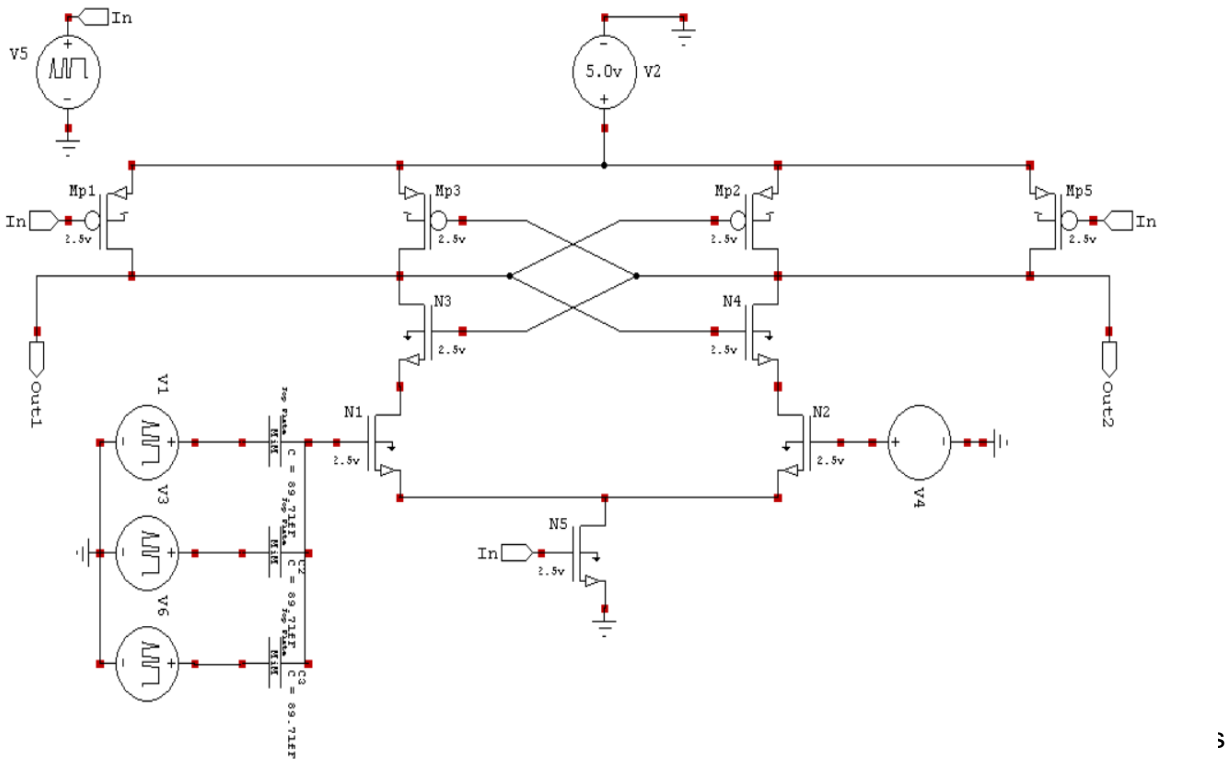
3. ONE CLOCK 3 INPUTS CCCL

One clock 3 input CCCL logic circuit is shown in figure 2.

In this circuit more than two inputs are used (three inputs). The circuit works in similar fashion as that of two inputs. The equivalent voltage from three input serves as input voltage in gate terminal of N1 MOSET and then it is compared with the provided reference voltage. The voltage comparison logic is same for three or more inputs as well which is shown in equations 1, 2 and 3.

In similar fashion n-input NAND/AND gate or NOR/OR gate can be designed just by connecting them with n number of capacitors to gate terminal of N1 NMOS and by selecting appropriate reference voltages. For example, if three input logic is to be constructed, depending on the input combination the node voltage of N1 will change. On setting appropriate reference voltage or the threshold value of threshold gate the desired output and complement of that can be obtained.

This is particularly a big advantage of using capacitor coupling dynamic logic over static logic. It also provides high speed and less static power.



NAND Fig. 2. Three input CCCL gate

4. SIMULATION RESULTS

4.1 OR/NOR Gate

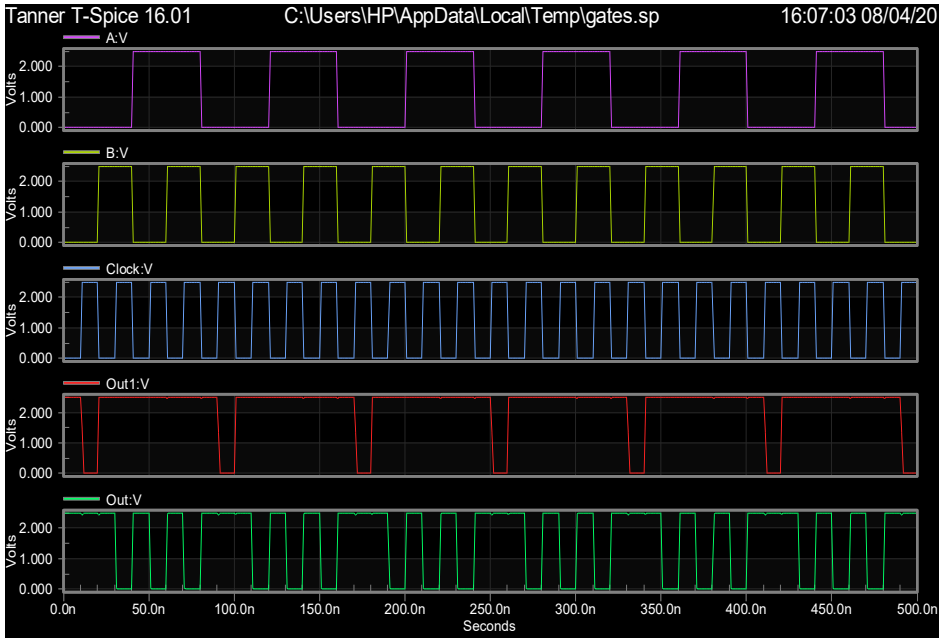


Fig.3. Output for OR/NOR gate

Figure3 shows the output of OR/NOR logic, When clk=1 and V_{REF}=0.75V and C_A and C_B are taken as 89.71fF.

By analyzing any point in the graph of out1, it satisfies OR output logic with respect to two inputs when clock is 1. Similarly it satisfies NOR output logic with respect to two inputs in the out2 graph.

4.2 AND/NAND Gate

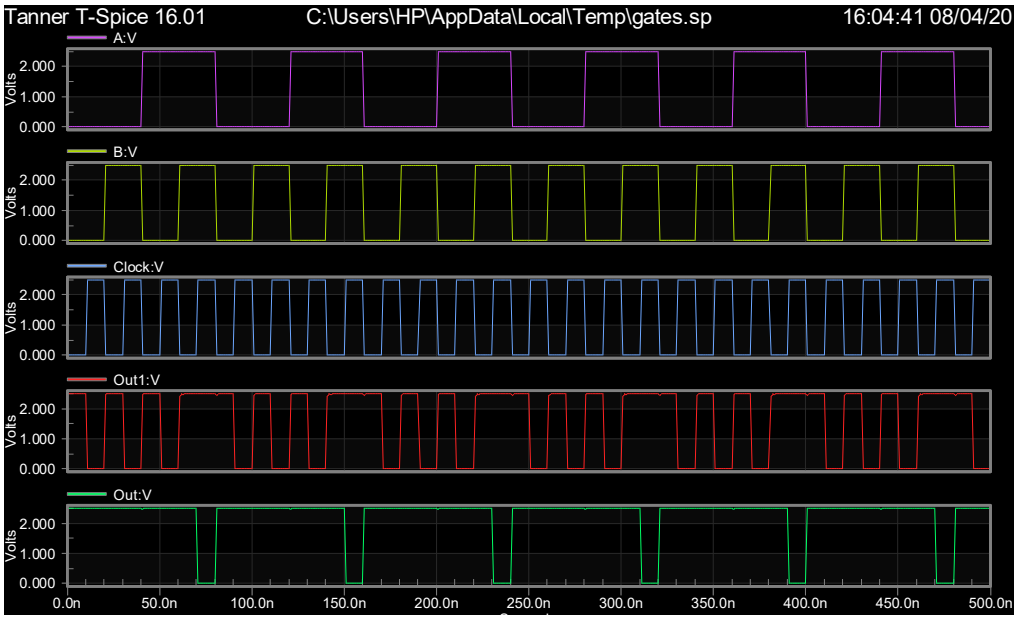


Fig.4. Output for AND/NAND

Figure 4 shows the outputs of AND/NAND logic, in out1 and out 2, when clk=1, V_{REF} =2.5V and C_A and C_B are taken as 89.71fF.

4.3 Majority/Majority NOT Gate

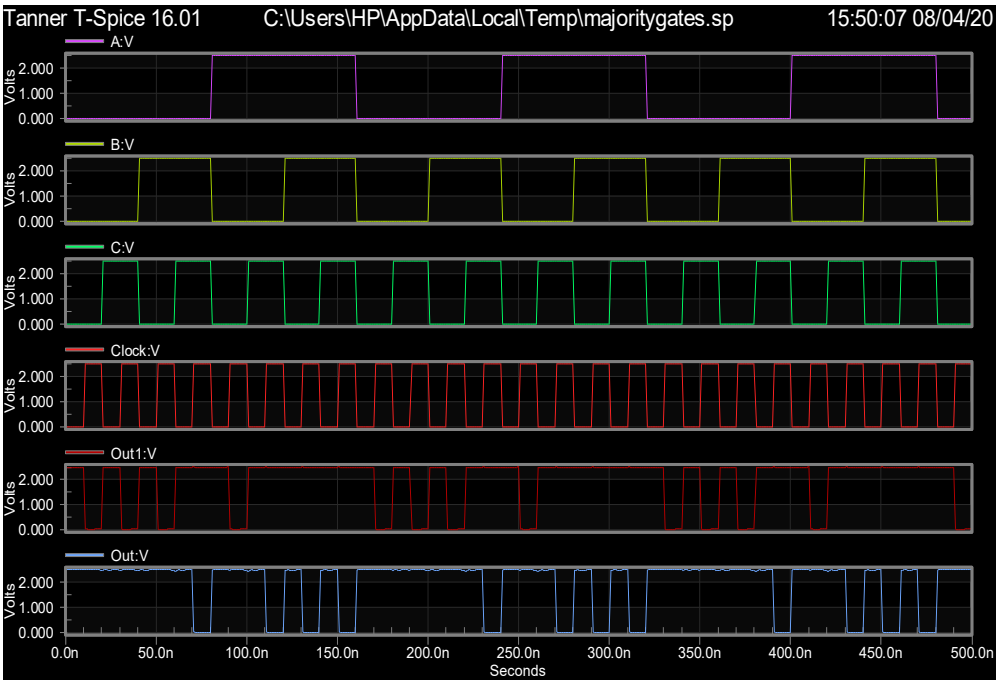


Fig 5. Output for Majority/Majority NOT Gate.

Figure 5 shows the outputs of Majority and majority NOT Gate logic where Vref is 1.25V. Hence, from the simulation studies, it is clearly observed that just by changing Vref set at various levels OR/NOR, AND/NAND & Majority/Majority NOT Gate operation can be easily obtained.

5. CONCLUSION

The multiple-input C3L based NOR, NAND, AND, OR, MAJORITY gate and MAJORITY NOT have been designed here using one capacitor coupling logic circuit. The simulation studies have been carried out using TANNER (T spice) software at 250nm technology node. Here, only one clock is used (instead of two clocks) as the reference voltage is directly applied without using any extra dummy and coupling capacitor. Less number of PMOS are used in this design. The smaller number of clock input substantially reduced the power consumption to $6.24*10^{-5}$ W for 20ns clock time period. The circuit proposed and tested here, is quite simpler with respect to the previously known design configurations. The unique technique, reported here, helps us to achieve low power and high-speed Threshold logic gates. The majority gate designed here has various applications in complex combinational circuits like Full adders. The work may be extended further to use single DC voltage source by generating a precisely variable reference voltage.

REFERENCES

[1] S. Vrudhula , N. Kulkami and J. Yang, "Design of threshold logic gates using emerging devices," 2015 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 373-376.

[2] M. Sarkar, G. S. Taki, Prerna, R. Sengupta and S. N. Ray, "Design of ripple carry adder using CMOS output wired logic based majority gate" 2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON), Bangkok, 2017, pp. 328-331.

[3] H. Ozdemir, A. Kepkep, B. Pamir, Y. Leblebici and U. Cilingiroglu, "A capacitive threshold-logic gate," in IEEE Journal of Solid-State Circuits, vol. 31, no. 8, pp. 1141-1150, Aug. 1996.

[4] L. Gao, F. Alibart and D. B. Strukov, "Programmable CMOS/Memristor Threshold Logic," in IEEE Transactions on Nanotechnology, vol. 12, no. 2, pp. 115-119, March 2013.

[5] C. Jia, L. Milor and H. Huang, “Capacitor coupling threshold logic," 45th Midwest Symposium on Circuits and Systems, 2002.

[6] Rouald H. Cieslak, “MOS Full Adder Circuit,“ U.S. Patent No. 4583192, Apr. 15, 1986.

[7] Hong-Yi Huang Teng-NengWang **,”**CMOS Capacitor Coupling Logic (C3L) Circuits” Proceedings of Second IEEE Asia Pacific Conference on ASICs. AP-ASIC 2000 (Cat. No.00EX434), Aug. 30, 2000.

[8] [1] Pui Ying Or, Ka Nang Leung, "A slew-rate enhancement technique based on current comparator and capacitive-coupled push-pull output stage for CMOS amplifiers", in IEEE Int. Conf. on Electron Devices and Solid-State Circuits, Hong Kong, EDSSC 2008, 2008, pp. 1-4.

[9] B. Razavi, "Design of analog CMOS integrated circuits", McGraw-Hill, New York, USA, 2001.

[10]Eslamifarsani, Majid & Ghaderi, Noushin. (2015). Reduction of coupling capacitance, using a capacitor-transistor coupling circuit. 10.1109/ELECO.2015.7394649.

[11] T. Gabara and W. Fischer, “Capacitive coupling and quantized feedback applied to conventional CMOS technology,” in CICC, San Diego, CA, May 1996, pp. 281–284.

A Review of Data Privacy Techniques for Wireless Body Area Networks in Telemedicine

Viktoriya Gorbach
School of Professional Study
New York University
New York, USA
vg1058@nyu.edu

Md Liakat Ali
Department of Computer Science & Physics
Rider University
NJ 08648, USA
mdali@rider.edu

Kutub Thakur
Professional Security Studies
New Jersey City University
NJ 07305, USA
kthakur@njcu.edu

***Abstract*—Wireless body area networks have become closely associated with the ever-evolving field of telemedicine. This paper provides a comprehensive review of the current applications of wireless body area networks (WBANs) in telemedicine, as well as the existing security measures in place to protect the sensitive patient data being stored and shared via potentially unsecured networks. Various types of proposed solutions will be reviewed, based on which a recommendation of the most effective solution will be.**

Keywords— WBANs, telemedicine, remote monitoring, data privacy, security, threats

I. Introduction

With the rapid growth of telemedicine over the last decade, there have been quite a few emerging technologies adapted for use in the field of medicine. One of these is the "the wearable". Sensor technology – through use of wearable biosensors (in devices such as the FitBit and Apple Watch) have been providing useful wellness and physical activity data to users directly, but in order to store and transfer this data to an alternate location, there needs to be a specialized network and communication architecture in place to do so.

Wireless body area networks (WBAN's), otherwise known as body sensor networks, are wireless sensor networks which are specifically designed to assist in connecting various types of health sensors and medical appliances through a series of nodes and actuators - to the internet. They consist of WBAN's, together with biosensor technology, have become an integral tool in remotely assessing, diagnosing, and monitoring a patient's health, especially in elderly populations and/or rural locations where there exists a known barrier to healthcare access. The implications of telemedicine on lowering the cost of healthcare for patients is certainly another driver for utilization. While there exist many non-medical applications for WBAN's and biosensors, this paper will focus on telemedicine and the protection of private health information (PHI) in the use of WBAN's. A number of research [1-9] have conducted surveys on WBAN's challenges and applications. Ensuring that the data transmitted via WBAN's is done so accurately is crucial for effective monitoring and response, however, protecting the integrity of a patient's private health information is equally as important. The rest of the paper is organized as follows: section II presents some WBAN sensor application, section III discusses WBAN architecture, section IV presents security/privacy requirements, section V presents existing solutions, and section VI is the conclusion.

II. WBAN Sensor Applications

A. Sensor Classification

There are three types of sensors nodes. Each of these is defined based on the way it is implemented. Below are the classifications:

1) Wearable: Sensor that is either placed on the surface of the skin or placed up to 2cm away from the body. Most medical applications of WBAN sensors utilize the body surface node. Examples are: EKG, EEG, blood pressure, heart rate, pulse oximetry, motion, sleep, and many others.

2) Implant: This type of sensor is planted either underneath the skin or inside the tissue of the body. Today, these types of sensors are most commonly used in diabetic patients for glucose monitoring, as well as in cardiovascular patients for purposes of monitoring for signs of myocardial infarction.

3) External: This type of sensor does not come in contact with the skin. Distance for implementation can range from several centimeters to five meters away from the human body. Mostly, these nodes are used for remote monitoring of a patient's vitals, or with the help of internet connectivity – for sending patient vitals and health status data directly to providers. This then, allows providers to provide remote assistance by video or phone call, for example.

The applications of WBAN sensors can vary greatly based on the medical specialty, purpose, and specific patient needs. Each application requires a different data rate, frequency, and consumption of energy. Table 1 [1] shows an example of some common WBAN sensor applications, and their individual specifications.

B. Types of Nodes

The different nodes in a WBAN network are categorized into three different types, based on their role within the network.

1) Coordinator: Acts as a gateway to a different WBAN, a trust center, or an access coordinator.

2) End: Does not have ability for data transmission to other nodes. It's restricted to perform a specific set of functions.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

TABLE I. COMMON APPLICATIONS OF WBAN SENSORS [1]

Application Type	Sensor Node	Data rate	Power Consumption	QoS (Sensitivity to Latency)	Privacy
In-body Applications	Glucose Sensor	Few Kbps	Extremely Low	Yes	High
	Pacemaker Endoscope	Few Kbps	Low	Yes	High
	Capsule	>2 Mbps	Low	Yes	Medium
On-body medical Applications	ECG	3 Kbps	Low	Yes	High
	SpO_2	32 bps	Low	Yes	High
	Blood Pressure	<10bps	High	Yes	High
On-body non-medical Application	Music for Headset	1.4 Mbps	Relatively High	Yes	Low
	Forgotten Things Monitor	256 Kbps	Low	No	Low
	Social Networking	<200 Kbps	Low	Low	High
Off-body Applications	Motion Sensor	35 Kbps	-	-	Low

3) Relay: Represents intermediate nodes, Consists of a parent node, child node, and relay messages. These nodes can sense data from other nodes.

Depending on the type and purpose of sensor, there may be an actuator capability present. Each of the above types of nodes provides information to the actuator mechanisms within a WBAN. The actuator is responsible for responding appropriately to the provided information (based on pre-written instructions). Actuator mechanisms can be the sole responding capability within a system, or one of many.

For example, in an implanted insulin pump – sensors within the tissue obtain a glucose reading, send the information to the actuator, which responds by administering supporting and appropriate amounts of insulin to a diabetic patient.

III. WBAN ARCHITECTURE

The basic architecture of most WBAN's will be discussed in this section. Along with describing the communication structure, the process of data transmission will be discussed here.

A. Communication Structure

1) Tier-1Intra-WBAN: At this level, the sensors are confined to communication within the body. The signals utilize a Personnel Server (PS) as a gateway to send the data to an access point (in Tier 2).

2) Tier-2 Inter-WBAN: It is on this level that access points are used to connect the information in a WBAN with other outside networks. There are two types: Infrastructure based and Ad-hoc based. The most commonly used architecture in WBAN's is infrastructure-based. This model allows for greater control over security, and more central management. In comparison, the ad-hoc based approach is larger and is able to extend coverage from two to 100m (allowing for movement through larger spaces).

3) Tier-3 Beyond-WBAN: This level is most suitable for medical databases, as it is here that the specific user information and medical history of each patient is stored. This level acts as a "gateway", connecting the specific WBAN to a larger metro area.

Tier 1 is where raw data is generated from the patient's wearable sensors, and from where this data is then distributed and stored. Tier 2 is where the data goes to be processed and aggregated before being stored again. The jump from Tier 2 to Tier 3 is done via Access Point, and at Tier 3 the data is sent to a much larger database/server for storage and access by authorized personnel (i.e: medical providers, insurance companies, scientists). Ensuring that this information goes from Tier 1 to Tier 3 and remains secure not only in route, but also for duration of storage in the central database is one of the greatest challenges of WBAN implementation [2].

B. Data Transmission

Data can be transmitted either from the device to the coordinator, or from the coordinator to the device. Specifically, in star technology, there are two methods of communication: beacon and non-beacon mode. Beacon mode is when the network coordinator is responsible for all communication by sending periodic 'beacons' to identify the start/end of a super-frame (which is set by the WBAN standard and the user).

In non-beacon mode, the opposite is true. A node is able to send data to the coordinator, but only when the coordinator is invited to participate in a communication. That is when they power up and ask the coordinator to receive data [2]. It's also very important to consider the utilized methods of data communication and transmission in evaluating the possibility of a data breach. Without understanding the existing architecture, it would not be possible to ascertain the limitations and/or address the problems.

For example, in the event of a node failure during transmission, it is crucial that the data remains unable to be

maliciously modified and its integrity/dependability remain intact. However, in order to protect this data, one needs to be aware of the involved components and flow of data/storage to implement safeguards at the appropriate places.

IV. Current Security/Privacy Requirements

After toolbar Data security refers to the safe transmission of patient data, while privacy refers to use exclusively by those intended to have access. The CIA triad (confidentiality, integrity, and availability) is quite relevant in the discussion of WBAN data security. One of the greatest challenges of data security and privacy in WBAN is achieving an optimal balance between availability, efficiency, and security.

The National Institute of Standards and Technology (NIST), is the federal agency which oversees the development of information security guidelines. To further support the work being done by NIST, there are several existing federal laws mandating adherence to the set forward guidelines.

The Health Insurance Portability and Accountability Act (HIPAA) was passed by Congress in 1996. Prior to enactment of HIPAA in 1996, there were no generally accepted guidelines for protecting health information in existence. In 2003, HIPAA came out with the Privacy Rule and the Security Rule. Each of these rules outlines very specific standards for privacy and security of a patient's sensitive health information, accordingly. Failure to adhere to these standards can result in fines and penalties for non-compliance. The Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009, which was enacted as a part of the American Recovery and Reinvestment Act, expanded the responsibilities of medical personnel and business associates to strengthen their protections of the privacy and security of patient health information.

With the increasing prevalence of e-health, telemedicine, and tele-monitoring – WBAN infrastructure has also evolved. And while the availability of complex biosensor nodes, wireless monitoring systems, and databases has grown, the availability of protection from malfunction or breach has not grown nearly as much as necessary. The first part of this section will address the most prevalent and common security threats to a WBAN system that exist today. The second part of this section will talk about the security measures currently in place to counteract such threats.

Table 2 provides an overview of common WBAN data security threats, as well as the associated security requirements. In addition, this table provides some of their proposed security solutions (which will be discussed in greater detail in the next section)

TABLE II. Common WBAN threats, requirements, and Solutions [2]

Security threats	**Security requirements**	**Possible security solutions**
Unauthorized access	Key establishment and trust setup	Random key distribution and Public key cryptography
Message disclosure	Confidentiality and privacy	Link/network layer encryption and Access control
Message modification	Integrity and authenticity	Link/network layer encryption and Access control
Denial of Service	Availability	Intrusion detection systems and redundancy
Compromised node	Resilience to node compromise	Inconsistency detection and node revocation and Tamper
Routing attacks	Secure routing	Secure routing protocols
Intrusions and malicious activities	Secure group management, Intrusion detection Systems and secure data aggregation	Secure group communication, Intrusion detection systems

A. Existing Security Threats

According to Xu et al [10], the attack model can help to categorize various types of security threats, according to the location of the attacker and type of attack in question. The authors define the external and internal attacks. External attacks refer to an attacker that sits among the nodes and gains sensitive data through the link layer. They state that the way to deal with these types of attacks is to employ cryptography and data disturbance technologies. Internal attacks refer to an attacker that obtains information and/or configuration parameters by specifically attacking a node, after which more private data can be obtained. According to the authors, the way to deal with this sort of attack is to use data anonymization and fuzzy techniques. Xu et al. also discuss attacking of eavesdropping private data versus attacking of tampering and forging private data. The former attack model infers data through analysis and monitoring of network traffic (which destroys the privacy and confidentiality of data in question), while the latter attacks nodes by means of tampering, forging, or deleting sensitive health information (which destroys privacy, authenticity, integrity, reliability, and availability of WBAN data) [10].

After reviewing several other papers and articles outlining the biggest security threats to the data in WBAN's, below are some of the most commonly discussed threats, in order of greatest frequency of mention [2, 10, 11]:

1) Node compromise/vulnerability: If a node is compromised, then it can no longer be trusted. Generally, this occurs once an attacker has found a node, and directly connected it to their computer. Once this happens, the attacker can proceed to utilize the node for their benefit and/or to harm others. A vulnerable node refers to a node that is not sufficiently protected from a data breach, therefore, posing a significant risk with each utilization for data collection/transmission.

2) Unauthorized access: Entry to (or use by) an unintended party of all of the nodes in the WBAN by pretending to be an authentic remote user. This can be a result of irresponsible behavior by intended users (eg: password sharing, etc) with no intention of malicious activity by the original intended user. This can also be a result of lacking network security. Regardless

of intent, this can lead to compromise of the nodes (eg: port scanning), the network (eg: installation of malware), along with breached integrity of all the data contained.

3) Denial of service (DOS)attacks: By utilizing a DoS attack on a WBAN, the network becomes unable to provide the necessary services. It is highly likely that any prolonged disruption of service will either lead to packet loss, failed transmission of vital patient data, or even a delay in appropriate and necessary life-saving intervention.

4) Jamming/tapering: Jamming is a specific subset of DoS attacks, in which radio frequency interference is utilized on several of the nodes to block the entire network. Jamming does not work on large networks, but works very efficiently on smaller networks. Since the current state of wireless links/nodes is generally highly exposed, jamming attacks are a high threat to WBAN networks, and can also lead to packet loss. Tampering refers to electronic interference, damage, or replacement of original information by an attacker. This could be done on a WBAN to obtain a patient's health information.

5) Intrusion/malicious activity: Refers to any number of types of attacks that a WBAN can be a subject of, which are of a malicious intent. Any activity that results in the inappropriate use or sharing of a patient's private health information would fall under this category.

B. Current Security Requirement

Most of the available literature differentiated between the most basic concepts (CIA triad and several others) that are required of a WBAN to keep data secure, and those advanced measures that are ensuring to maintain security levels above the current standard (even still with some limitations and lack of certainty). Below are listed some of the basic measure concepts:

1) Data confidentiality: Ensuring that data remains unexposed and protected from unauthorized access.

2) Data integrity: Refers to measures taken to ensure the accuracy and consistency of content; this applies both to singular data as well as data packets.

3) Availability of network: Crucial aspect of WBAN; the network should be available at all times to the patient, and in case of loss of availability – the network should be able to seamlessly change operations to an alternate WBAN.

4) Data authentication: Allows nodes within the WBAN to be verified as originating from a known trust center or trustworthy node, by accurate calculation of a Message Authentication Code (MAC).

5) Secure management: Coordinator should have secure control, in order to deliver key distributions via decryption/encryption techniques.

6) Dependability: Failure in retrieving correct patient data, and in a timely manner, can be life-threatening.

7) Locating abilities: Ability of a WBAN to be able to correctly identify a patient's location in order to prevent an attacker from being able to send fake location signals during an attack, for example.

8) Accountability/flexibility: This refers to both the necessity of authorized providers to be held responsible for upholding their duty in keeping sensitive PHI secure, as well as to the patient's right to have ability to change access controls of their WBAN in case of emergency or change of provider.

9) Privacy rules/compliance requirements: HIPAA, HITECH (discussed in more detail at the beginning of this section).

Following are some of the most commonly mentioned advanced security methods being employed today, with discussion of their limitations [1,2, 12].

1) TinySec: Came about as a solution to gain link layer encryption and authentication of data. The process of the system is to employ a group key between sensor nodes, while encrypted data packets and a MAC are calculated for the entire packet. Limitation: relies on a single, shared key – which provides minimal security against physical node capture.

2) Biometrics: The intrinsic characteristics of a patient's body are used to manage cryptographic keys for sensors attached to the body. In other words, two of the specific sensors need to measure an equal value so that a symmetric key can be securely generated and distributed. Limitation: complex in nature, and difficult to maintain when mobile patients connected to one WBAN reach the proximity of another WBAN.

3) IEEE 802.15.6 security protocol: The most updated 2012 version of the standard, which strives to provide an international norm for short-range wireless communication around a human body, which utilizes reliable low power, can range anywhere from narrow to ultra-wide band (75.9Kbps to 15.6Mbps). Includes two layers of specifications:Physical and MAC layer. Physical layer deals with Narrowband (NB), Ultra Wideband (UWB), and Human Body Communications (HBC). This layer defines the electrical and physical specifications. MAC (Media Access Control) layer deals with control of the hardware responsible for interaction with the wireless transmission medium.

4) ZigBee security services/IEEE802.15.4: In addition to the outdated IEEE protocol, ZigBee provides supplementary security protocol/standard. It does this by identifying a trust center, through which coordinator nodes become responsible for allowing other nodes to join the network and to appropriately distribute keys. It addresses the unique needs of low-power, wireless IoT networks, precisely like the WBAN. Limitation: Low complexity and low data speed of the ZigBee protocol, as well as its high cost of maintenance surely place Zigbee (and the outdated IEEE 802.15.4 protocol) at a disadvantage when considering future solutions.

5) Bluetooth/Bluetooth Low Energy(BLE): Bluetooth is designed for short range communications with high data rates. Operates in 2.4GHz frequency and works similarly to a star topology. But it has some limitations such as it consumes high energy and is not suitable with a WBAN application, as they are already low-energy networks. On the other hand, Bluetooth Low Energy (BLE), also known as Bluetooth 4.0, was designed

to operate on lower-power and in shorter range. In WBAN applications, BLE can pair devices and synchronize in milliseconds. The limitations of BLE are similar to those of traditional Bluetooth technology, however, some mobility support for WBAN applications exists.

6) Wireless security protocols: Wi-Fi Protected Access (WPA) ad Wi-Fi Protected Access version 2 (WPA-2) are two of the protocols being used for wireless network security. WPA uses temporal key integrity protocol (TKIP) for encryptions, while the newest WPA-2 version uses AES, which is much more secure and reliable. Limitation: Not able to cover all traffic, and not able to protect the nodes.

7) Encryption technologies: Hardware; Elliptic curve cryptography; Symmetric Key Encryption: Algorithm that utilizes the same key for encryption of plain text and decryption of cipher-text; Conventional Public Key Encryption: Technique that utilizes a paired public and private key algorithm to secure data communication; Identity-based Encryption: One of the most primitive types of cryptography. It uses some unique information about the identity of the user, as the user's public key.

V. Solutions

The following are some of the most commonly proposed solutions for minimizing the potential of a data breach at any point in the WBAN, as per the literature review. Each of these are either only theoretical concepts just yet or are still in the early stages of implementation for WBAN systems. Consequently, there are pros and cons to each of these solutions, as well as some impacts that are as of yet unknown.

A. Software Defined Networking (SDN)

Offers a technological approach to network management that could improve performance. It would centralize network intelligence by separating the control plane from the data plane. It is expected to reduce the complexity of network management by enabling dynamic and efficient network configuration. In WBAN implementation, this would render the network more like that of a cloud-based system than the current and traditional methods. Also, it has the potential for a great reduction in the node's energy consumption, but not enough research has been done on this yet to ensure this outcome. Fig. 1 depicts the architecture of an SDN.

Limitation: It is not yet known how SDN will function with heterogeneous (and other) data aggregation, which may either become a pro or con of SDN utilization. Many of the other theorized benefits have yet to be explored and finalized, which makes the effective implementation of SDN a limitation for the time being [1, 13].

B. Blockchain

Emerging blockchain management systems can offer increased security of the highly sensitive obtained information by creating an effective and secured data-sharing platform. Blockchain is essentially a decentralized and distributed ledger which stores digital information in a public database.

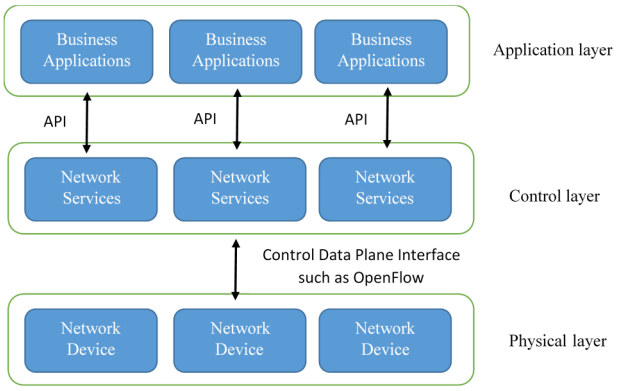


Fig. 1. SDN system architecture

Limitation: Modification of data in blockchain is made impossible by its immutability property. This leaves the potential for erroneous data to be entered into a patient's electronic health record (EHR) if a dysfunction occurs in the network.

C. Intrusion Detection and Prevention Systems (IDPS)

Can be directly used for WBAN network protection, but more so discussed in terms of a digital forensics tool for potential discovery of malicious activities, and to aid in development of improved security/privacy policies for WBAN servers (as well as other m-health applications and databases.

D. Energy Harvesting (EH)

Energy harvesting is the 'harvesting' of energy from external sources. This energy is then captured, stored, and utilized. EH provides a small amount of power to low-energy electronics. Mostly used for small, wireless networks precisely like the WBAN, it becomes a self-sustainable energy source for the network.

Limitation: Does not generate sufficient energy quickly enough for most applications, and the modules required tend to be bulky.

E. Multi-biometric scheme

Multi-level approach to security, which essentially utilizes a specific part of the body/node sensor to authenticate the identity of an authorized/intended user and provides a 'key'. When this is utilized in conjunction with other authentication and biometric-based methods, it becomes a multi-biometric scheme. Khan et. Al. [14], proposed a cloud-based framework for improved data security, and while this may very well be a viable option for the future of mobile healthcare applications, it is more interesting that their approach also utilized a multi-biometric system in order to generate a common key.

Limitation: Can at times be dependent on the environment of the biosensors. For example, proper functioning may not occur at a certain room temperature or light exposure.

F. Advanced Encryption Standard (AES)

While there are already some cryptographic methods being utilized in WBAN to ensure the integrity and confidentiality of

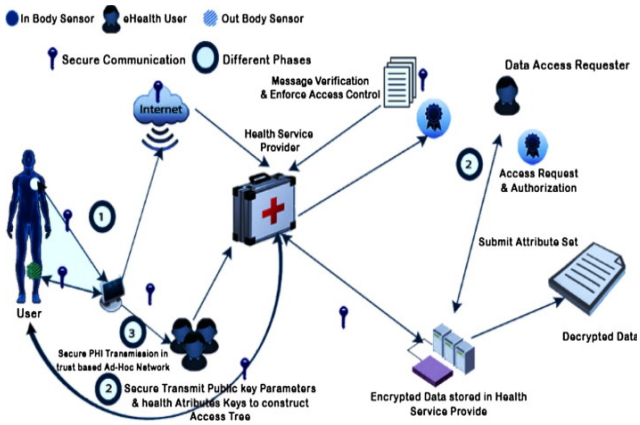


Fig. 2. WBAN security and privacy [2]

data transmission by encrypting the bio-signals, employing a 256-bit AES will decrease the likelihood of a data breach even further. Established by the NIST, the original encryption standards were created in order to provide specifications and requirements for the encryption of electronic data [15]. Figure 2 depicts a typical example of security and privacy in a WBAN [2]. An appropriate location for most of the above-mentioned solutions could be identified according to this scheme as it is, but some of the solutions would require an altered scheme in which they could be included.

VI. Conclusion

When it comes to WBAN technology and telemedicine applications, data security and patient privacy are directly correlated. If the collected information is not able to be confidentially and securely transmitted to the intended provider and/or healthcare facility, then it is no longer reliable. In a worst-case scenario, if such a breach results in interrupted transmission or falsified data, then the result could prove to be life-threatening or fatal for the patient involved.

After conducting a literature review, a pattern began to emerge regarding potential methods of enhancing the security of patient data in WBAN transmission and storage. Based on existing evidence and theories, the research recommends for increasing the breadth and depth of data security in WBAN would be a hybrid approach of existing solutions. The combination of SDN and a multi-biometric scheme would provide greater security both for the physical and software sides of the network. It would increase security of patient monitoring, ensure efficient and secure delivery of collected data, support energy efficiency of the nodes, and provide a higher percentage of secure keys available on the network. This hybrid approach to solving the issue of data security in WBAN's is likely to yield the highest increase in data security/privacy of the network, while requiring the least amount of initial investment.

In the future, however, much more resource needs to be done on the SDN in order to be able to fully understand its capabilities and implications for WBAN implementation, as well as to improve on its current limitations and unknown factors. Also, more resources should be invested into further improving the multi-biometric security approach. While overall the method is straight-forward, the specifics of WBAN implementation still need to be fine-tuned in order to be easily applicable to all types of WBAN networks and biosensor nodes uniformly.

References

[1] K. Hassan, K. Biwas, K Ahmed, N. Nafi, K. Islam, "A comprehensive review of wireless body area network," *Journal of Network and Computer Applications*, vol. 143, pp 178-198, 2019

[2] S. Al-Janabi, I. Al-Shourbaji, M. Shojafar, S. Shamshirband, "Survey of main challenges (security and privacy) in wireless body area networks for healthcare applications," *Egyptian Informatics Journal*, vol. 18, pp 113-122, 2017.

[3] S. Akbari, "Energy harvesting for wireless sensor networks review." In *2014 Federated Conference on Computer Science and Information Systems*, pp. 987-992. IEEE, 2014.

[4] M. Chen, S. Gonzalez, A. Vasilakos, H. Cao, and V. C.M. Leung. "Body area networks: A survey." *Mobile networks and applications* 16, no. 2 (2011): 171-193.

[5] G. V. Crosby, T. Ghosh, R. Murimi, and C. A. Chin. "Wireless body area networks for healthcare: A survey." *International Journal of Ad Hoc, Sensor & Ubiquitous Computing* 3, no. 3 (2012): 1.

[6] F. Akhtar, and M. H. Rehmani. "Energy harvesting for self-sustainable wireless body area networks." *IT Professional* 19, no. 2 (2017): 32-40.

[7] Thakur, Kutub, Meikang Qiu, Keke Gai, and Md Liakat Ali. "An investigation on cyber security threats and security models." In *2015 IEEE 2nd International Conference on Cyber Security and Cloud Computing*, pp. 307-311. IEEE, 2015

[8] R. Cavallari, F. Martelli, R. Rosini, C. Buratti, and R. Verdone. "A survey on wireless body area networks: Technologies and design challenges." *IEEE Communications Surveys & Tutorials* 16, no. 3 (2014): 1635-1657.

[9] S. Pathania, and N. Bilandi. "Security issues in wireless body area network." *Int J Comput Sci Mobile Comput* 3, no. 4 (2014): 1171-8.

[10] G. Xu, Q. Wu, M. Daneshmand, Y. Liu, M. Wang, "A data privacy protective mechanism for wireless body area networks," *Wireless Communications and Mobile Computing*, vol. 16, pp. 1746-1758, 2015.

[11] C. Jang, D. Lee, J. Han, "A Proposal of Security Framework for Wireless Body Area Network," *2008 International Conference on Security Technology*, Hainan Island, pp. 202-205, 2008.

[12] M. Anwar, A. Hanan Abdullah, R. A. Butt, M. W. Ashraf, K. N. Qureshi, and F. Ullah. "Securing data communication in wireless body area networks using digital signatures." *Technical Journal* 23, no. 02 (2018): 50-55.

[13] W. Lee, Y. Choi, N. Kim, "Study on Virtual Service Chain for Secure Software-Defined Networking," *Advanced Science and Technology Letters*, vol.29, pp. 177-180, 2013.

[14] F.A. Khan, A. Ali, H. Abbas, N.A. Haldar, "A cloud-based healthcare framework for security and patients' data privacy using wireless body area networks," *Procedia Computer Science,* vol. 34, pp. 511-517, 2014.

[15] R. Khalilian, A. Rezai, E. Abedini, "An Efficient Method to Improve WBAN security," *Advanced Science and Technology Letters*, vol.64, pp. 43-46, 2014

Potential Development on Cyberattack and Prospect Analysis for Cybersecurity

Lippter Li
School of Professional Study
New York University
New York, USA
pl1962@nyu.edu

Kutub Thakur
Professional Security Studies
New Jersey City University
NJ 07305, USA
kthakur@njcu.edu

Md Liakat Ali
Department of Computer Science & Physics
Rider University
NJ 08648, USA
mdali@rider.edu

***Abstract*—Technology and methods on cyberattack and strategy on cybersecurity always develop simultaneously. Once a knotty method of cyberattack has been created, information security specialists always react immediately and come up with a new security system that can prevent and eliminate the attack. Then, the hacker will try to find bugs from the system and perform alternative attacks. The war on security just never ends. The article tries to conclude our current information security environment and frontier battleground. The article is focused on some specific techniques such as machine learning and neural networks along with their applications on both the cyberattack side and cybersecurity side. The future predictions are made on how A.I. technology may change the current situation on information security. How attackers may use these technologies to steal user privacy and how a current security system should adapt to the evolution of cyberattack. Google has announced the realization of quantum supremacy on a quantum computer. As known, a quantum computer can solve problems required years with a current computer in seconds. This can be catastrophic to modern cryptology. The article discusses the problem and introduces potential reconstruction solutions on encryption. The ultimate goal of the article is to discuss the potentials with modern technology on information security and alarm people to be aware of these risks, also making predictions on how the future of information security will be like.**

Keywords— Cyber Security, Cyberattack, User's Privacy, Security System, Information technology, Machine Learning, Neural Networks, Artificial Intelligence, Quantum Supremacy

I. INTRODUCTION

Information security is one of the most heavily invested areas by the government and corporations nowadays. This is due to the number of massively increased crimes on cyberattacks. In the first six months of 2017, the recorded number of data stolen cases exceeded two billion, which is already double the number reported of the whole year 2016. Two billion dollars' ransomware payments were also announced, making the situation more severe than ever [1].

A. History of Information Security

Ever since the early days of communication, the war between message keepers and message breakers never ends. In the ancient era, due to lack of highly efficient communication methods, important military instructions were usually sent by messengers or ravens. However, without cryptography, such delivery can easily be stopped, and the messages were known ahead by the enemy. Around 50 B.C., the great commander Julius Caesar created probably the first encryption methodology called "Caesar Cipher". Since then, the main focus on information security is on how to create a more complex encryption system. However, for more than a thousand years, the method of procedural handling controls remained the same. Messages were kept transported by trusted persons, guarded and stored in a secure environment or sturdy boxes.

In the 19th century, the invention of telegrams changed the total gaming rule in information security. Now people have a method to exchange messages without physical letters. In 1861 to 1865, the American Civil War was the first warfare that used telegrams as a method of transport messages. Warring parties use this technology to manage military deployment and give commands. Telegram used radio as a medium of transportation. This gives the opposite party a chance to spy the message. Soon, the method of monitoring the enemy message became realized; the telegram was no longer a safe choice for message communications.

Entering the 21st century, the popularity of the interconnected network on computers makes another way of exchanging private messages possible. The Internet uses the internet protocol suite (TCP/IP) to connect the computer device, creating a convenient communication system throughout the world. People who use a computer can use the internet to send messages to another computer all around the world as long as the device is connected to the internet as well. The delivery of an important message has never been this easy before, however, the disclosure of messages has never been this easier as well.

B. Current Status of Information Security

In ancient times, the transportation of secret messages relies on couriers. When telegram was invented, you needed to be near the receiving center to spy incoming mails or to enter the information center to steal the hardcopy message. As a result, physical security always attracted more attention back then. CCTV and levelled authorization systems were created to keep the suspicious away from the information center; however, with the creation of the internet, there now exists the possibility of hacking into the data center while the hacker is thousands of miles away. In 2013, a couple of American Banks including Chase, Citi Bank, JPMorgan suffered DDoS attacks

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

simultaneously after investigation, the attackers were located in Iran which is over 10,000 kilometers from the States [2].

On the other hand, the variety of different cyberattack methods make it more challenging to perform protection as well. DDoS, phishing, SQL injection and even social engineering are all potential methods used by hackers. These make the current information security situation more severe than ever.

C. Current and Potential Methods to Eliminate Cyberattacks

Thankfully, our modern technology used by information security specialists is always one step ahead of the hackers'. Researches and data have shown more than 92% of cyberattacks can be prevented by a well-designed security system, reliable operational performance and timely incident handling. In fact, around 86% of cybercrimes were finally uncovered, and the criminals are punished.

What's more, while more and more investments are running into the information security area, it already becomes the test field of new technology. Machine Learning has already been proven extremely efficient when dealing with DDoS attacks; neural networks and other A.I. methods start to be widely used on security structure and system implementation. Hopefully one of the first application areas of quantum computers will be on information security as well. We will discuss each of the above new solutions and their perspective below.

II. Machine Learning

A. Introduction to Machine Learning

Machine Learning is a statistical method as well as a computational algorithm firstly announced by American computer scientist and statistician Arthur Samuel in 1959. With more than half-century of development, machine learning has developed to a complete theory of data analysis and classification methodology.

According to Mitchell [3], the underlying ideology of machine learning is to let the computer program acquire knowledge from observation with respect to some class of tasks. In order to achieve the goal, the program will firstly separate existing data sets into two separate groups. One is called the training set; the other is called the testing set.

The software will first operate on the training dataset. It performs universal modification and injects each data point onto the hyperplane, then draws the best fit separation line on that hyperplane to perform clustering or to forecast potential patterns shown in figure 1.

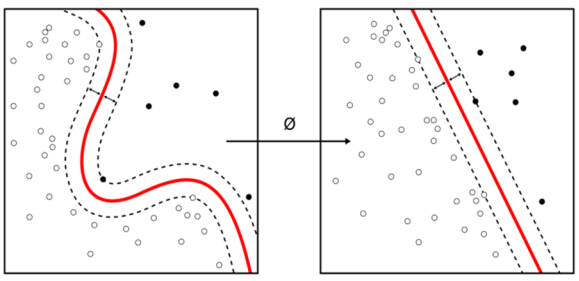


Fig. 1. Best fit separation line [3]

After the clustering is finished, the software uses the testing dataset to re-test the pre-established model. The reason for these steps is that in real-life, outliers and extreme cases may occur randomly. These cases may profoundly influence the final model; therefore, cause inaccuracy. On the other hand, the procedure is also used to delete irrelevant factors from the dataset to create a more realistic model. The software uses the testing dataset to test the accuracy of the model and perform minor changes to prevent over-fitting.

Typically, machine learning is separate into three approaches to conquer different queries of problems.

1. Supervised learning has input value and returns cost function to measure the validity of the model. The algorithm has clearly target output returns. Successfully implemented supervised machine learning models can make close output estimations or correct classifications.
2. Unsupervised learning has no desired output. It usually deals with examples with unknown patterns and without pre-existing labels. The software requires to figure out the rule all by itself. Clustering problems are usually solved by this method. The software finds similarities and dissimilarities between each example and gives the optimal clustering group to separate data.
3. Reinforcement learning exists somewhere in between, instead of a clear output target. The algorithm only has a vague reward mechanism that gives reward on some specific behaviors. The software's target is to find the optimal strategies to maximize the reward. Reinforcement learning is widely used in a lot of game simulations and puzzle-solving.

B. Machine Learning in Cyberattack

Machine learning is a very efficient tool to manipulate. Tasks that take for hours or days can be easily shortened to seconds with the application of machine learning software. The situation is also true when performing cyberattacks.

Phishing is the most common way of cyberattacks. The attacker disguises himself as a trustworthy entity in electronic communication and then verbally deceives critical personal data such as credit card information or implements worms or viruses on the victim's computer via suspicious files. Within the first three months of 2017, more than 100,000 cases of phishing crime were reported in the United States. Some of the infamous enterprise and political scandals such as Target data breach and Hillary Clinton email controversy. Phishing is so prevalent among hackers because it is an extremely low cost and unbelievable easy way of attack. Viruses and worms are very simple to code or can easily be downloaded from the internet. The only mission attackers required to do is to spread the message out to as many people as possible and expect some of them to open the file. As a result, the most difficult part of performing a phishing attack is to trick the victim into believing the hacker's fake identity.

As mentioned above, Hackers usually pretend to be some official entity or close friends. In order to gain beliefs from the victim, the hacker will try to gather some personal information prior to the attack to increase the success rate. Machine learning,

as a wonderful data analytics tool, has a spectacular performance on public personal information collection and analysis. In the current information era, everyone has some personal information exposed to the public on social media. Machine learning can be used to collect this information and with only the victim's name as a key word. The model can self-select and clustering all information that is actually related to you can eliminate the others. Machine learning techniques can be used to classify all your information such as your closest friend, home address, telecom vendor and family relationship. After that, it can further analyze your characteristics. The model can predict your future movement, potential requirements and even personalities. Be aware, all these collections, classification, and analysis all happened in seconds with the help of machine learning algorithms. By far, the hacker has already controlled your detailed profile which significantly increases the possibility to trick you. In fact, hackers usually use machine learning for characterization even before the victim selection. Hackers, after having collected thousands of emails, will send malware only to those who would click on the link. The name-list of who is easy to get tricked by is all provided by machine learning analysis.

C. *Machine Learning in Cyber Security*

Machine learning itself is not a bad technology that only brings negatives. Actually, it brings more positives than negatives to the cybersecurity area. Machine learning performing surprisingly well on prevent and handling DDoS attacks which is one of the most troublesome cyberattacks that bothered security specialists for years [4-5].

DDoS stands for Distributed Denial of Service shown in figure 2 is a cyberattack that the hacker floods the target machine or server with superfluous requests in an attempt to overload the system. The most challenging part of dealing with DDoS attacks is that in distributed denial of service attacks, the incoming traffic flooding originates from different sources. Some of the IP addresses are real IPs that are controlled and stolen by the attacker, and some of the IPs are simply just fake; However, it is impossible to monitor which of these requests come from the attacker and which of them actually come from real clients since every request has different IP addresses. Before the invention of machine learning, the most common way to deal with DDoS attacks is simply shut down the server entirely so that no request at all will be processed [6].

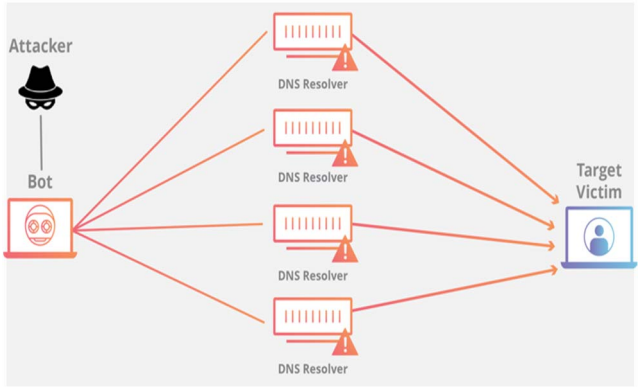


Fig. 2. Example of a Distributed Denial of Service [8]

Recent researches have shown that machine learning has incredible behavior while solving DDoS attacks. Even though the IP addresses may seem like a series of random numbers to us, it is full of logic and patterns for computers. With machine learning algorithms, it is actually possible to find common patterns in network packets that carry out the attack and therefore eliminate most of the attacker's fake IP addresses all at once. The Catholic University of Korea's professionals, Noh et al. [8], has run tests on their security program with supervised machine learning algorithms. It turns out that the software has the ability to recognize 87% of the fake addresses with the correct rate of 90%. What was more surprising was, after detailed analysis, the software can actually recognize a hacker's potential real IP address and therefore, locate the hacker's physical position, which makes arresting possible. Besides, some IP address patterns are often used by hackers to perform DDoS attacks. The result can be simply calculated by the software, and put those addresses on the blacklist. As a result, every time those addresses submit some request, the system will be alarmed or the system can simply refuse to receive requests from the address any more.

The result is phenomenal. DDoS is definitely the most troublesome cyberattack. The cost and inconvenience of servers shut down brings colossal damage to real clients and corporations every year. With this application of machine learning, enterprises no longer need to be afraid of DDoS attacks in the future. The method of preventing flooding of hacker's requests while still keeping the server functional is possible.

III. Applications of Neural Networks in Cyber Security

A. *Introduction to Neural Networks*

The idea of using electronic circuits to imitate the human brain's functionality was long proposed by Scottish psychologist Alexander Bain. However, the ideology failed to come true until Canadian computer scientist Geoffrey Hinton came up with his updated methodology and application called backpropagation a hundred years later in 1985.

Backpropagation used a mathematical optimization method called gradient descent as a support. Gradient descent is based on the observation on general multivariable functions that if the function (F(x)) is defined and differentiable in a neighborhood of a point α, then F (α) decreases the fastest in the direction of the negative gradient of F. Therefore, we can have the equation:

$$\alpha_{n+1} = \alpha_n - \gamma \cdot \nabla F(\alpha_n) \quad (1)$$

γ is a suitable step size.

Backpropagation used this ideology to create multilayer web-like networks to simulate the function of neutrons. The first layer is called a input layer, where it stores every data collected. The last layer is called the output layer where it returns the final output. The middle ground is called hidden layer, where calculation and modification are performed. The model has a loss function which can be used to calculate the accuracy of the final output. Basically, the algorithm starts with the random arrangement, then compares the final result with the actual value to calculate differences. After that, gradient descent is used to backwardly distribute and revaluate changes and functions in the

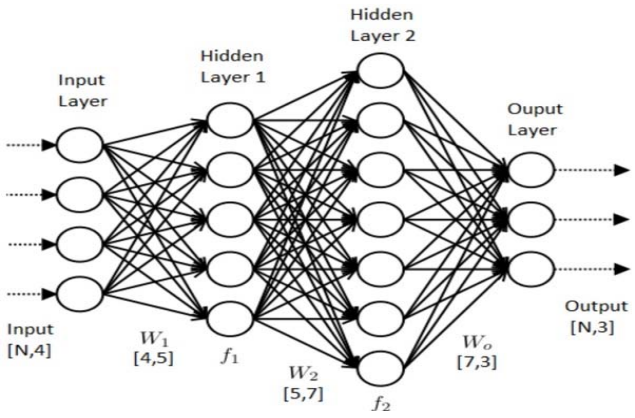


Fig. 3. Neural Network with 2 hidden layer [8]

hidden layer. The process keeps on iterating until the result calculated is exceptionally close to reality. Consequently, we achieved our final model of the problem. The process is shown in figure 3 [8, 9].

B. *Neural Networks in Cyberattacks*

Compared to standard machine learning models, neural networks models do not show superiority in accuracy. In fact, compared to traditional machine learning, the accuracy of the neural network actually decreases. The purpose of neural networks is trying to let computers think as humans. Humans make mistakes, as neural networks. As a result, the model by neural networks may not be extremely accurate; however, it is definitely the most human-like.

Hackers can use this quality to perform low-cost cyberattacks. Hackers often use neural networks to create communication robots to reply messages. Reports have shown that the modern phishing and ransomware crimes no longer rely on single mailing communication. Since people nowadays are familiar with the crime, internet users are careful when receiving suspicious emails. Consequently, instead of sending out the viruses and worms immediately, these hackers use auto-reply robots to pretend as the receiver's friends or family members, so that they relax email receivers in the first place. As the conversation goes further, the hacker then sends out the viruses which gigantically increase the success rate. These robots can analyze the receiver's personal relationship on social platforms, then choose a fake identity from the receiver's contact list and start the lively conversation with the victim. This new form of attack is proved 5 times more effective than traditional methods. Even some information security professionals fall into the trap. Sometimes, the hacker is not even required to inject viruses or worms. Simply pretending to be the victim's friends and asking for a money transfer can also succeed. What's more threatening is that a single computer performs all these procedures. What's more, this computer can reply and analyze hundreds of victims at the same time, create unbelievable efficiency on frauds.

C. *Neural Networks in Cyber Security*

However, computers are still computers, with careful operations, it is still easy to prevent becoming the victim. Neural networks bring more contribution to the information security area than damages. The modern information security system architecture is very dependent on the technology. Neural networks are instrumental in prevention and defending new malware invasion.

Malware is a type of harmful software designed to cause damage to the injected computer. Common types of malware including viruses, worms and Trojan horses. A computer virus is a piece of code that modifies another computer program to cause dysfunctionality when activated. A worm is a type of software that when executed, will continuously replicate itself until the server is full and crashes. Trojan horse is a combination of both virus and worms as well as a backdoor software that provides unauthorized access to the controller.

The structure of these malwares is surprisingly simple. Some of them may only contain less than 100 rows of codes and compare to the malware twenty years ago. The coding ideology hardly changed. However, computers nowadays can still easily get infected by these viruses because even though the security system is familiar with every pre-established malware, tiny minor modification is all required to build a new kind of virus. Therefore, even though the functions are the same, the system is unable to recognize and identify the software and have no prepared solution for the attack as well.

Neural networks, since simulate real human brains processes, can deal with unfamiliar attacks. While foreign new malware is injected, the neural network model can quickly analyze the danger level of the software to determine whether the incoming software is malware or not. Even if the software fails to detect the danger of the software, After the software starts to function. The model is possible to create suitable anti-virus software to stop the attack on time. In addition, the software actually can perform self-testing practice. Developing itself by first creating a new kind of malware and then trying to eliminate the malware to progress its functionality. In the near future, neural networks will be the best malware defending protector as well as the best malware creator.

D. *Other Artificial Intelligence Applications*

Both machine learning and neural networks are actually different approaches for artificial intelligence in solving different categories of problems. Of course, there are lots of other approaches which are very useful in information security as well.

Symbolic artificial intelligence is one of the oldest ideologist of A.I. Symbolic A.I. solves problems with logic and mathematical computation. Symbolic A.I. machines start with basic rules and information (data), then use logical thinking and statistical analysis to revert the thinking process and draw conclusions according to the calculation. The method remains high accuracy on solving mathematical and logical problems which in the information security area, to create or decipher passwords. People tend to create passwords that are hard for other to guess but easy for themselves to remember. The brutal method of guessing passwords is trying every possible combination of the letters which may take years to finish. However, with the help of A.I. technology, the machine can analyze a user's public information and simulate the user's original logical thinking process, then come up with a group of possible passwords based only on public information. This may

sound horrible; however, the same methodology can also be used to create passwords. The machine can help in creating the least likely password but at the same time, easy to remember [10].

Evolutionary algorithms are another different approach to artificial intelligence. Instead of simulating the human brain, the algorithm chooses to imitate the evolution processes. The algorithm first creates several potential solutions, then chooses the best fit for reproduction, and breeds new individuals through crossover and mutation operations. After generations of reproduction, the final model is created. As a result, the method is very handy to create lethal malware, as well as impenetrable firewalls [11].

IV. Quantum Computer and Quantum Supremacy

In 1941, the creation of Atanasoff-berry Computer announced the first automatic electronic digital computer. Since then, people have been using electronic signals and binary systems to operate computers. In the 1980s, along with the development of quantum physics, physicist Paul Benioff suggested using quantum particles, instead of electronic particles, to create a computer. This is the first recorded annunciation of quantum computers. Paul Benioff suggested using quantum particles instead of electrons to create the computer due to the theory of superposition principle.

We choose binary coding in computer science because there exist only two states of electrons, positive and negative. However, in quantum physics, the state is more complicated. Quantum superposition states that any two quantum states can be added together, and conversely, any quantum state can be represented by the addition of two quantum states. As a result, in quantum physics, the signal can be positive, negative, as well as positive and negative at the same time. Imagine a 4-bit memory traditional computer meaning that the computer can only state four-bit data (example of data includes 1001, 1100, 1111, 0000); as a result, in order for all combinations to be present, 16 of this kind computer is required. However, with a 4-bit memory quantum computer, since the state in a quantum computer is very vague, assuming all quantum bits are in superposition, only one computer is required to present all combinations. So, as in extreme cases, one quantum computer has the ability to compete for 16 traditional computers if the size of memory is restricted to 4-bit [12].

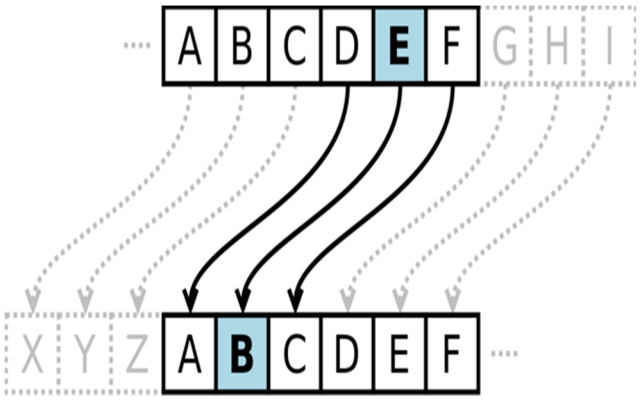


Fig. 4. Example of Caesar cipher [13]

In fact, assuming the traditional computer has the memory of size n, the potential of same size memory quantum computer equals 2^n of traditional computers all combined The power of processing and ability to solve problems increases exponentially as the memory increases. This potential that quantum computers can ultimately exceed the power of traditional computers and even solve problems that are impossible for electronic computers is called quantum supremacy. With the processing and calculation power of the quantum computer, even solving the question in the most brutal way, a quantum computer can still defeat the electronic computer in speed. Problems that used to require years to solve can be shorted to seconds. This is catastrophic to modern cryptography.

A. History of Cryptography

The first encryption techniques were created by Julius Caesar around 50 B.C. to prevent leakage of important military messages. Caesar cipher is an elementary type of substitution cipher shown in fig. 4. The methodology is merely replacing each letter of the message with some fixed number of positions down the alphabet. For example, with one number of positions down the alphabet, the word "*apple*" can be encrypted to "*bqqmf*". As time passes, the complexity of encryption becomes higher and higher. During world war II, German created a special encryption machine called "*Enigma*" to be responsible for encryption of important telegram messages. Special decryption team full of mathematicians and cryptography professionals were formed by the Allies to solve this mystery. It took these genius years to come up with the solution finally. The successful decryption was announced one of the vital factors of the victory in World War II. In 1977, with the popularity of computers. American cryptographer Ronald Rivest and his teammates created the best known and popularly used numerical encryption method called RSA [13, 14].

RSA (Rivest-Shamir-Adleman) named after its inventors, is a public-key cryptosystem used in the modern information security area. The ideology is based on the mathematical theory that the level of difficulty of factorization increase as the number increases. In other words, it is literally impossible to factorize a large number that is the product of two substantial prime number. Assume Tom want to send and receive messages from his friends via an untrustworthy tunnel. In order to prevent spying. Tom firstly find two different large prime number, p and q, let $N = p \cdot q$ be the product of these two prime number. Then, let $r = (p-1)(q-1)$. Choose an integer $e < r$ such that e and r are relatively prime. Then they can easily achieve d where $e \cdot d \equiv 1 \ (mod\ r)$. Finally, the combination (N, e) is called to be the public key that can show to everyone who wants to send message to Tom, and (N, d) is called private key that needs to keep in private and use only on decryption. Assume Tom's friend, Jerry, receives Tom's public key and want to send message to Tom, he first tries to change his message into a series of number that both he and Tom can understand. Then, he decrypts the message by calculating $c \equiv n^e \ (mod\ N)$where n is the message and e, N are Tom's public key. Jerry will send only the number c to Tom. After receiving Jerry's message, Tom can decrypt the message by calculating $n \equiv c^d \ (mod\ N)$ where n is the message and d is the private key. Apparently, with the above procedure, the outcome source can't decrypt the message unless he knows the private key d; however, to acquire the private key,

the hacker needs to solve the factorization problem on giant number *N* which is the product of two large prime numbers. In fact, when the number is large enough, the process may take hundreds of years to solve with our current computer's processing power. As a result, RSA is the most used encryption method in current society, especially in finance area. Almost every bank in the world uses RSA to performing currency exchanges between customers. However, even though the brutal force is proved impossible to break RSA, with Quantum Computer, it is another story.

B. *Cryptography and Quantum Computing*

The processing power of quantum computers is proved to be supremacy. Factorization of large numbers may be unsolvable for current computers, but not for quantum computers for sure. Modern encryption methods like RSA, will be extremely fragile. However, it is stupid to worry about our privacy's safety. Quantum computing will bring a more secure encryption method that keeps our information safer than ever. Quantum cryptography is the science of exploiting quantum mechanical properties to perform cryptographic tasks. It solves the crisis and chaos quantum computing may bring to cryptography and creates a new system of encryption. There exist many theories on how to use quantum computing to encrypt and secure our system. One of the potential methods is called quantum key distribution which uses similar ideology to RSA [15].

Assume Tom wants to communicate with his friends while preventing being spied by the third party. They will create a pair of shared random secret keys used for both encryption and decryption. Since the secret key is in a quantum state, any measurement by the third party will introduce detectable anomalies and change the outcome of the key. As a result, both the receiver and sender will sense the spying immediately. The method is actually more secure than RSA, since it not only perfectly keeps the message in private, but also has the alarming system that monitor any incoming suspicious and when spying occurs, it reminds both sides immediately. Important advantage of this quantum key distribution (QKD) method is that instead of using unsolved mathematical theory, QKD uses physical reality to create the system which makes the system more reliable.

V. CONCLUSION

People need to be aware of the new technologies. Sometimes, these technologies may bring trouble to our daily lives. The machine learning algorithm can be used to collect our personal information; neural networks can be used to perform social engineering; artificial intelligence can calculate your passwords and create the most potent malware; with the application of quantum computers, all current encryption methods will become meaningless. On the other hand, Machine learning is proved to be super effective on DDoS attacks; neural networks can be used to create the most powerful firewall; quantum computer, while making all current encryption methods functionless, also established a new, more secure cryptography system.

The rivalry between cyberattack hackers and cybersecurity protectors will never end. In some ways, this rivalry is healthy to the entire society. Technologies that are created and used in this massive attack and protect war are gradually promoting the whole human society as well. Machine Learning is also trained to help operating high accuracy surgeries and save thousands of lives every year. Neural Networks is the primary method used on autopilot vehicles. Quantum computers have the potential and ability to let our human find solutions to the unsolvable and reach the area we never touched before.

Overall, technology is neutral. No technology is meant to harm. Only when technology fails to be controlled by the right hand will the tragedy come true. We need to be aware and caution with the development of technology. We need to get familiar with how these technologies may affect us in negative ways and be prepared ahead for them. At the same time, embrace the positive and live happier than yesterday.

REFERENCES

[1] M. Fosco, 'Will artificial intelligence save us from the next cyber attack?' Ozy (30 Oct 2018) www.ozy.com/fast-forward/will-artificial-intelligence-save-us-from-the-next-cyber attack/88428/, Accessed April 2020

[2] E. Jardine. "Global cyberspace is safer than you think: real trends in cybercrime." (2015).

[3] T. Mitchell, *Machine Learning*. McGraw Hill. p. 2. ISBN 978-0-07-042807-2, (1997).

[4] G. Kumar, K. Thakur, and M. R. Ayyagari. "MLEsIDSs: machine learning-based ensembles for intrusion detection systems—a review." *The Journal of Supercomputing* (2020): 1-34.

[5] A. Polyakov, "Machine Learning for Cybercriminals." *Medium*, Towards Data Science, 31 July 2019, towardsdatascience.com/machine-learning-for-cybercriminals-a46798a8c268, Accessed April 2020

[6] Ali, Md Liakat, Kutub Thakur, and Beatrice Atobatele. "Challenges of Cyber Security and the Emerging Trends." In *Proceedings of the 2019 ACM International Symposium on Blockchain and Secure Critical Infrastructure*, pp. 107-112. 2019.

[7] Brun, Olivier, Yonghua Yin, and Erol Gelenbe. "Deep learning with dense random neural network for detecting attacks against iot-connected home environments." *Procedia computer science* 134 (2018): 458-463.

[8] S. Noh, C. Lee, K. Choi, and G. Jung. "Detecting distributed denial of service (ddos) attacks through inductive learning." In International Conference on Intelligent Data Engineering and Automated Learning, pp. 286-295. Springer, Berlin, Heidelberg, 2003.

[9] Ali, Md Liakat, Kutub Thakur, and Charles C. Tappert. "User Authentication and Identification Using Neural Network." *i-manager's Journal on Pattern Recognition* 2, no. 2 (2015): 34.

[10] M. Buscema, "A Brief Introduction to Artificial Neural Networks." *Artificial Adaptive Systems in Medicine: New Theories and Models for New Applications* (2009): 5.

[11] Thakur, Kutub, Meikang Qiu, Keke Gai, and Md Liakat Ali. "An investigation on cyber security threats and security models." In *2015 IEEE 2nd International Conference on Cyber Security and Cloud Computing*, pp. 307-311. IEEE, 2015.

[12] P. Benioff, "The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines." *Journal of statistical physics* 22, no. 5 (1980): 563-591.

[13] C. M. Bishop. *Pattern recognition and machine learning*. springer, ISBN 978-0-387-31073-2, 2006.

[14] D. S. Berman, A. L. Buczak, J. S. Chavis, and C. L. Corbett. "A survey of deep learning methods for cyber security." *Information* 10, no. 4 (2019): 122.

[15] B. Schneier, "Quantum cryptography: As awesome as it is pointless." *Wired, October* (2008).

Enabling event-related potential assessments using low-density electrode arrays: A new technique for denoising individual channel EEG data

Sujoy Ghosh Hajra
School of Engineering Science,
Simon Fraser University

Flight Research Laboratory,
Aerospace Research Centre,
National Research Council
Canada
Ottawa, Canada
sujoy.ghoshhajra@nrc.ca

Shishir Gopinath
School of Computing Science,
Simon Fraser University
Burnaby, Canada
shishir_gopinath@sfu.ca

Careesa C. Liu
School of Engineering Science,
Simon Fraser University

Rotman Research Institute,
Baycrest Health Sciences
Toronto, Canada
cliu@research.baycrest.org

Gabriela Pawlowski
Department of Biomedical
Physiology and Kinesiology,
Simon Fraser University
Burnaby, Canada
gmp5@sfu.ca

Shaun D. Fickling
School of Engineering Science,
Simon Fraser University
Burnaby, Canada
sficklin@sfu.ca

Xiaowei Song
Health Science and Innovation,
Surrey Memorial Hospital,
Fraser Health Authority
Surrey, Canada
xiaowei.song@fraserhealth.ca

Ryan C.N. D'Arcy
School of Engineering Science,
Simon Fraser University
Burnaby, Canada
rdarcy@sfu.ca

***Abstract*— Background: Brain function assessments based on event-related potentials (ERPs) derived from electroencephalography (EEG) are increasingly being conducted in realistic out-of-the-laboratory settings for clinical and non-clinical uses. For rapid testing and practical limitations, such applications require the use of low-density electrode arrays. A major impediment to their use in these applications is the lack of denoising techniques capable of removing artefactual contamination and isolating the ERPs features of interest within low-density arrays. Methods: A novel denoising technique combining empirical mode decomposition (EMD) with template matching procedure is developed and applied to individual-channel data, and the results of this new approach are compared to the results of a conventional (independent component analysis) denoising approach. Both whole-epoch morphological comparisons and specific ERP feature amplitude comparisons were undertaken at the group and individual level for a variety of ERPs indexing sensory (N100), attention (P300) and language processing (N400) using data from 31 healthy adults. Results: The new denoising technique successfully enables the capture of ERPs ranging from low-level sensation to attention to language processing (all $p<0.05$). Intra-class correlation analysis reveals high degree of similarity in the time series waveforms derived from the new and the conventional denoising approaches for all ERPs (highest $r=0.89$, all $p<0.001$). Analysis of specific ERP features of interest reveals no significant differences between the ERP amplitudes of the waveforms generated using the two techniques, and Pearson correlation suggests a high degree of similarity at the individual level (0.88 for N100, 0.78 for P300, and 0.80 for N400, all $p<0.05$). Conclusion: The new denoising technique is capable of operating on individual-channel EEG data, and produces results that are similar to those produced by conventional denoising techniques that use data from large whole-head electrode arrays. This new approach may thus enable more widespread use of ERP techniques in real world settings with low-density electrode arrays.**

Keywords*— *biomedical signal processing, neurotechnology, electroencephalography (EEG), neural signal processing, empirical mode decomposition (EMD), artifact removal

I. Introduction

Electroencephalography (EEG) is a non-invasive technique for capturing the the electrical activity of the brain through the use of scalp electrodes [1]. EEG-derived event-related potentials (ERPs) index a variety of cognitive processes, and have shown promise in both diagnostic and prognostic applications for various brain disorders [2], [3]. Cumbersome equipment and long testing times have traditionally confined ERP techniques to specialized research laboratories [4]. However, in recent years, the emergence of portable EEG hardware has led to increasing interest in applying ERP-based techniques in more realistic environments (e.g. patient bedside). Recently developed rapid automated assessment protocols, such as brain vital signs and Halifax Consciousness Scanner (HCS), have solved the challenge of rapid testing compared to the hours traditionally needed for ERP assessments [2], [5]. Similarly, the emergence of miniaturized and portable EEG hardware with low-density EEG sensor arrays has provided the requisite technology

978-1-7281-9615-2/20/$31.00 ©2020 IEEE ©2020 National Research Council Canada

platforms to rapidly collect EEG data. However, a key hindrance to the actual use of low-density EEG sensor arrays paired with rapid assessment protocols is the lack of techniques for denoising the EEG data for extraction of ERP signals.

ERPs are often embedded among larger (usually orders of magnitude) unrelated noise from background EEG and other physiological (e.g. ocular), instrumental, and environmental (e.g. mains power) sources [6]. While several techniques for denoising EEG data and extracting ERP signals of interest currently exist, these techniques are largely ill-suited for use with low-density EEG systems. As an example, the brain vital signs and HCS systems mentioned above utilize data from three EEG electrodes located at the top of the head. However, in order to reliably denoise the data using conventional EEG denoising approaches and isolate the ERP signals, they require two additional electrodes around the eyes - representing a 67% increase in the number of sensors/electrodes used. The need for such increased electrode set necessitates additional set up time which is an obstacle for use in clinical environments and also increases the resourcing requirements and cost of the hardware itself, hindering the use of truly low-density electrode arrays. A technique that is capable of operating in low-density electrode arrays – with the extreme case being 1 electrode – is thus needed.

Presently, many EEG data are commonly denoised using independent component analysis (ICA), a blind source separation based technique that enables the identification and removal of physiological and other noise through decomposition of multichannel EEG data into distinct component sources [7]. ICA is popular due to the technique requiring no prior knowledge of EEG and noise sources, but requires large electrode arrays for proper decomposition as the number of identified sources is limited by the number of electrodes. Recent extensions of the ICA technique have proposed using time-shifted segments of data to denoise individual electrode time series using multidimensional ICA [8]. However, a key requirement of this technique is that the underlying sources not have overlapping frequency content between desired signals and undesired noise [8]. Unfortunately, this criterion is not met in cases of most ERP studies (e.g. contamination from ocular sources has frequency content in the overlapping range as the neural signals of interest). Thus, both the traditional ICA approach (requiring large electrode arrays) and the newer individual electrode ICA extensions (requiring non-overlapping frequency content) are unsuitable for denoising ERP data in low-density electrode arrays.

Other approaches commonly used for removing noise and artefact from EEG data includes regression and adaptive-filtering based techniques. Examples of such approaches include the longstanding Gratton and Coles technique, which is a regression based approach, utilizing data from additional electrooccuogram (EOG) electrodes around the eyes along with the EEG channels, whereby the relative attenuation of the ocular artefact from the anterior to the posterior sections of the head is used to calculate factors to estimate and remove ocular contamination for each electrode [9]. Adaptive filtering techniques, on the other hand, utilize the EOG channel data as a reference to iteratively calculate filter weights necessary for the removal of the ocular artefact from the EEG channel data [10]. Since both of these techniques rely on using reference electrodes in the form of EOG to provide sufficient information about signals of ocular origin, they are ill suited for use within low-density electrode arrays as they necessitate the use of additional electrodes to capture the ocular artefact faithfully.

Similarly, other common approaches include the use of wavelet transform to decompose the individual electrode signal into component waveforms, or the use of principal component analysis (PCA) which transforms the data into N-dimension space by rotating the axes [11]. However, wavelet transforms can lead to smearing of temporal boundaries, especially in low frequencies [12], and PCA suffers from the same challenges as ICA in that it requires a large number of electrodes to sufficiently distinguish signal and noise sources. Other available approaches rely on the application of machine learning techniques [11]. However, such machine learning techniques are primarily used for artefact detection rather than for correcting the influence of the artefact in the data [11].

Another approach to denoising EEG data is through empirical mode decomposition (EMD), which a time domain signal decomposition approach ideally suited to non-stationary signals like EEG [11]. Several studies have investigated the application of EMD to various aspects of EEG processing: For instance, Williams et al. utilized EMD to isolate salient signal features in EEG data surrounding the 3Hz frequency [13], while Sweeney et al. used EMD combined with canonical component analysis to remove motion artifact from data [14]. However, given that ERP signals occur in a broad frequency range and are often contaminated by a variety of artifacts, these prior approaches could not enable adequate denoising of EEG data. In order to address the challenge of denoising EEG data within low-density electrode arrays, this study presents a novel technique using EMD combined with a stereotyped template matching (EMD-TM) procedure capable of denoising ERP data at the individual level. In particular, our technique uses template matching to retain salient signals of interest corresponding to ERP features while removing other components. This paper first presents the technical details of the new technique, then applies it to denoise individual-electrode data, and finally compares the performance of the new EMD-TM technique with that of gold-standard technique techniques such as ICA applied to the same EEG dataset.

II. Methods

A. Participants

Data from 31 healthy adults (aged 32 ±11yrs, 15 females) were collected as part of this study. Participants were right-handed, fluent in English, with normal hearing and no history of neurological issues or psychoactive medications. The research ethics boards at Simon Fraser University and Fraser Health Authority approved this study. Prior to data collection, all participants provided written informed consent.

B. Data Acquisition

Participants were presented with an auditory stimulus sequence consisting of interlaced tones and spoken word pairs, as described in detail in prior works [2], [15]. The tone stimuli consisted of 67% standard tones (75dB) and 33% deviant tones (100dB) and were designed to elicit neural markers corresponding to sensory (N100 ERP) and attentional (P300

ERP) processing. The word pairs consisted of congruent (e.g. bread-butter, 50%) and incongruent (e.g. bread-cat, 50%) pairs and were designed to elicit neural makers of language processing (N400 ERP).

All data were collected in a dedicated EEG suite with consistent conditions across participants. Auditory stimuli were presented binaurally using Presentation software (Version 18.0, Neurobehavioral Systems Inc., USA) and insert-type earphones. EEG data were collected using an EEG system with Ag/Ag-Cl electrodes (BrainAmp 64-channel with ActiCap). Electrode impedances were maintained at less than 20kOhms, and raw data were collected using BrainVision Recorder (Version 1.20.0801 Brain Products GmbH) and saved to file for processing.

C. Data Analysis

Raw EEG data were processed using two streams – one using the conventional ICA denoising and one using EMD-TM. The general workflows of both streams are shown in Fig. 1. In order to ensure that the two approaches were truly independent, separate researchers performed each stream blinded to the other stream.

D. ICA Stream

EEG data were downsampled to 500Hz, and inspected for noise and re-referenced to linked mastoid electrodes (TP9 and TP10) in line with previous literature [15]. Subsequently, a 0.1-50Hz zero-phase shift 4th order Butterworth bandpass filter was applied to the data and ICA was undertaken using the InfoMax algorithm to identify and remove components corresponding to ocular, cardiac, muscular and other noise sources based on temporal, spatial and spectral characteristics as described in previous works [16], [17]. ICA strives to resolve the observed activity into a set of independent source activity as described in equation (1):

$$\boldsymbol{s}(t) = W \cdot \boldsymbol{x}(t) \tag{1}$$

where *x(t)* are the observed/captured signals (dimensions sensor x time), *s(t)* are the estimated source signals (dimensions component x time), and W is the weight matrix to transform between the two spaces. In this study, the InfoMax technique was employed for estimating the weight matrix using a maximum likelihood approach, i.e., W is estimated in an iterative manner by following the natural gradient with learning rate μ as shown in equation (2):

$$\Delta W = \mu(I - 2\tanh(\boldsymbol{x})\,\boldsymbol{x}^T)W \tag{2}$$

Following denoising using ICA, the data were low-pass filtered (0.5-20Hz), segmented (-100 to 900ms relative to stimulus presentation), baseline corrected (-100 to 0ms relative to stimulus presentation) and conditionally averaged to generate ERPs corresponding to each stimulus type.

E. EMD Stream

Raw data were downsampled to 500Hz, and referenced to linked mastoid electrodes (TP9 and TP10). A 0.1-50Hz zero-phase shift 4th order Butterworth bandpass filter was applied to the data, and the data were segmented (-100 to 900ms relative to stimulus presentation), and EMD based decomposition was applied to each epoch of each stimulus type on a per-channel basis. EMD, a nonlinear signal processing technique, is considered to be well suited for non-stationary signals such as EEG and consists of decomposing the time series signal into a set of intrinsic mode functions (IMFs). The IMFs are generated using the following steps:

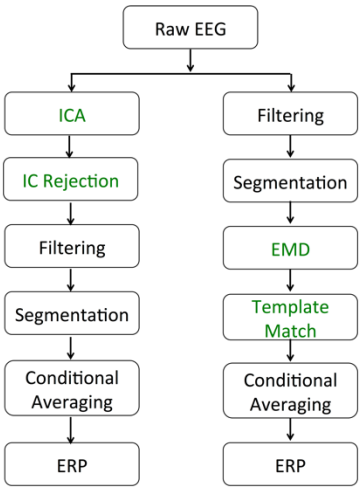


Fig. 1. Overview of processing streams. Major steps of the traditional ICA-based (left) and proposed EMD-based (right) pipelines undertaken as part of this study, with the key denoising steps highlighted in green for both streams of analysis.

1. Identify all the local extrema (minima and maxima) across the entire length of the time series s[n]
2. The identified maxima and minima are connected using separate cubic spline functions to create the upper, u[n], and lower, l[n], envelopes
3. The mean of the envelopes are calculated, m[n]= [u[n] + l[n]]/2
4. A proto-IMF is calculated as the difference between the data and the mean calculated in the step above, h[n] = s[n] – m[n]
5. Check if proto-IMF satisfies the conditions of IMF, i.e., *i*) the difference between number of maxima and the number of zero-crossings is zero or one over the entire length of the data, and *ii*) the mean value of the envelope defined by the maxima and the envelope defined by the minima must be zero.
6. If proto-IMF does not satisfy criterion in (5), then repeat above steps on h[n]
7. If proto-IMF satisfies criterion in (5), proto-IMF becomes IMF component, c[n]
8. Subtract the IMF from time series signal to create residual, q[n] = s[n] – c[n].
9. Repeat steps (A) to (H) on residual, q[n]. EMD operation concludes when the q[n] is a monotonic function.

The IMFs calculated using the above-mentioned technique, c_k, can be added together to reconstruct the original signal as shown in equation (3), where r_n is the residual after extraction of n IMFs:

This work was partly supported with funding from MITACS, and the Multi-Year Funding program at Simon Fraser University supported SGH while grant funding from CIHR supported CL.

$$s = \sum_{k=1}^{n} c_k + r_n \quad (3)$$

Once the IMFs were calculated for each epoch, the IMFs were subjected to a template-matching procedure to identify the components to retain and to remove. In particular, the template for each participant was generated by averaging together the ICA-cleaned data from all other participants excluding the current participant. For each IMF, the degree of correlation between the IMF and the template was calculated in the time window depending on the ERP of interest (as described below), and if the correlation passed a threshold it was marked for retention, otherwise it was marked for removal. The threshold value was calculated using equation (4) derived from the calculation of correlation coefficient and its significance for any two generic time series, with t specified as 1.96 to correspond to 95% confidence level:

$$threshold = \sqrt{\frac{t^2}{(stop-start+1)-2+t^2}} \quad (4)$$

where *start* and *stop* correspond to the beginning and end of the time windows of interest in the epoch corresponding to the ERP features (i.e. 50-450ms for isolating N100 and P300 ERP, and 125-625ms for isolating N400 ERP). All IMFs marked for retention were summed together to generate the cleaned ERP waveform. The entire process described above was repeated for each participant within a leave-one-out framework to create the template from ICA-cleaned data and generate denoised data as output.

F. Performance Evaluation

In order to assess the performance of EMD-TM compared to the more conventional ICA-based approach, focused evaluations were undertaken at three electrode locations at the top of the head along the midline encompassing the anterior-posterior axis. Data from three participants were removed from subsequent analysis since they did not have data from all three electrode locations.

G. Whole-Epoch Morphological Comparison

Morphological assessment of the whole epoch was undertaken through evaluation of consistency of ERP waveforms derived from the two denoising approaches using intra-class correlation (ICC) with norm-referenced reliability [12]. Specifically, this entailed calculating the ICC on individual-level trial-averaged waveforms derived using the ICA and EMD-TM denoising approaches, and subsequently averaging the ICC values across participants. Statistical significance was established using a variant of the Monte Carlo estimates, i.e., a bootstrapped distribution of ICC values was created by randomly permuting the waveforms across participants and the two denoising approaches and recalculating the ICC by averaging the subject-level ICC values. The permutation process was repeated 1000 times to generate a null distribution of ICC values against which the significance of the true ICC values could be compared, with $p<0.05$ considered significant.

H. Signal (ERP effect) Capture with EMD-TM

Since the goal of this study was to create a new technique for denoising data for ERP experiments, the ability of EMD-TM to denoise data and capture the ERP effects of interest was evaluated. ERPs are deflections in the time series waveform of specified time windows relative to stimulus presentation, and exhibit different polarities depending on experimental condition. For example, the P300 ERP is a positive deflection in the 250-500ms range presenting larger amplitude in the deviant stimulus condition relative to the standard stimulus condition. To assess the extent of signal capture, the mean amplitude of each ERP waveform within specific time windows of interest was calculated, and compared across pairs of experimental conditions. In particular, the N100 and P300 ERPs were computed within the 110-160ms and 225-400ms time windows, respectively, applied to both the standard and deviant experimental conditions and chosen based on previous literature [15]. Similarly, the N400 ERP was computed in the 375-575ms time windows and applied to the congruent and incongruent experimental conditions. Statistical comparisons were conducted using 2-way ANOVA (with 'electrode location' and 'stimulus condition' as factors) with Huynh-Feldt correction for violations of sphericity, followed by pairwise comparisons using paired T-tests with Bonferroni correction.

I. Comparison of Signal Captured by ICA vs. EMD-TM

To compare the performance of the EMD-TM technique with that of ICA, the mean amplitudes of ERP waveforms derived from the ICA-cleaned data were calculated using the same time windows and stimulus conditions as specified in the previous section. The mean amplitudes from the two techniques were statistically compared using a series of paired T-tests or Wilcoxon Signed Rank tests depending on normality of data, with Bonferroni correction for multiple comparisons. To further assess the ERP effects captured by the two denoising techniques, the mean ERP response amplitudes corresponding to the deviant tone and incongruent words from the two techniques were subjected to Pearson correlation. These conditions were chosen as they reflect the experiential manipulations that give rise of the N100, P300 (deviant tone) and N400 (incongruent word) ERPs.

III. Results

A. Whole-Epoch Morphological Comparison

Fig. 2 shows the grand-averaged ERP waveforms across all participants derived using both the traditional ICA pipeline and the new EMD-TM approach. Results show that EMD-TM and ICA both produced ERP time series waveforms that are very similar in morphology as well as amplitude of the ERP components (N100, P300 and N400). A more quantitative assessment was also undertaken using the ICC approach described previously, and the results are shown in Table I. As indicated in Table I, the new EMD-based processing pipelines generated ERP waveforms that were highly consistent with the conventional ICA-based processing pipeline (all $p<0.001$). Figure 3 shows the ERP time series for a representative participant.

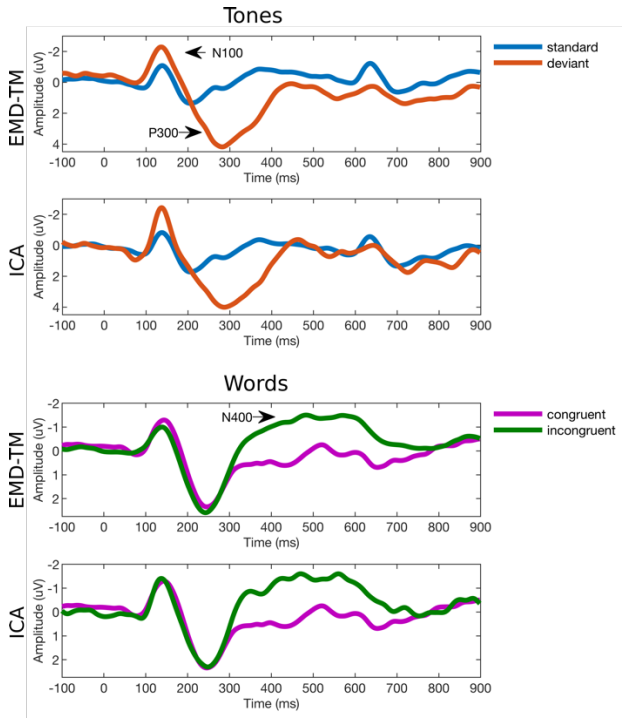


Fig. 2. Grand averaged ERP waveforms. Results shown for both EMD-TM and ICA approaches derived from the Cz electrode. Results shown for tones (N100 and P300 ERP) and words (N400 ERP), with components indicated using black arrows.

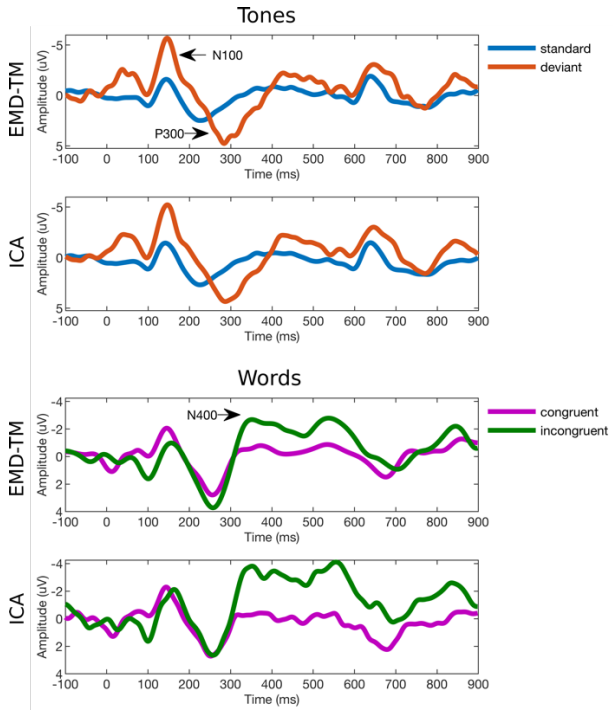


Fig. 3. Individual-level ERP waveforms. ERP time series for a representative subject at the Cz electrode are shown for tones (N100 and P300 ERP, top panel) and words (N400 ERP, bottom panel).

B. ERP Effect Signal Capture Using EMD-TM

The EMD-TM technique successfully captured the main experimental effect of stimulus condition for all ERPs (N100: $F(1,27)=32.9$, $p<0.001$; P300: $F(1,27)=87.1$, $p<0.001$; N400: $F(1,27)=50.7$, $p<0.001$), consistent with known characteristics of these ERP components [2, 15]. Additionally, the main effect of electrode location was also observed (N100: $F(1.72,46.4)=8.44$, $p=0.002$; P300: $F(1.38,37.3)=4.63$, $p=0.027$), which is also in line with known spatial features of these ERP components [15]. An interaction effect of stimulus condition and electrode location was also observed (N100: $F(2,54)=3.82$, $p=0.02$; P300: $F(1.63,44.0)=11.1$, $p<0.001$; N400: $F(2,54)=4.89$, $p=0.011$). An effect of electrode location was not observed for N400 ($F(1.47,39.6)=0.71$, $p=0.46$). As shown in Fig. 4 the ERP components isolated from waveforms derived using EMD-TM successfully captured the differential amplitudes expected across the experimental conditions at every electrode. The pairwise comparisons of mean amplitudes conducted on a per-electrode basis shows the increased negative amplitude for the deviant condition compared to the standard condition for N100, the increased positive amplitude for deviant relative to standard condition for P300, and the increased negative amplitude for congruent compared to the incongruent condition for N400 (all $p<0.05$).

TABLE I. ICC RESULTS FOR MORPHOLOGICAL COMPARISONS.

Stimulus Type	Frontal	Central	Parietal
Standard	0.79 ± 0.03*	0.89 ± 0.02*	0.80 ± 0.04*
Deviant	0.74 ± 0.05*	0.78 ± 0.05*	0.74 ± 0.04*
Congruent	0.64 ± 0.06*	0.72 ± 0.04*	0.66 ± 0.05*
Incongruent	0.67 ± 0.05*	0.82 ± 0.03*	0.72 ± 0.04*

Mean individual-level similarity between waveforms generated using the two denoising techniques. *p<0.001.

C. Comparison of Signal Captured by ICA vs. EMD-TM

As shown in Fig. 5, the mean amplitudes of ERPs derived from waveforms generated using the two denoising techniques did not reveal any significant differences. Indeed, as shown in Fig. 6, the ERP amplitudes generated by the two denoising approaches were significantly correlated (all $p<0.05$) – with the highest correlation coefficients of 0.88 for N100, 0.78 for P300 and 0.80 for N400.

IV. DISCUSSION

The goal of this current study was to generate a new technique for denoising individual-channel EEG data for ERP experiments. Towards this end, a novel technique was developed that combines EMD with a template matching procedure designed to retain stereotypical ERP response characteristics. The results indicate that the new EMD-TM technique produces waveforms that are substantially similar in morphology across the entire epoch compared to those generated by the traditional denoising (ICA-based) approach. Indeed, comparisons of the specific ERP amplitudes showed that the responses generated using the new and the traditional approach did not differ significantly from one another across a variety of ERPs indexing sensation (N100), attention (P300) and language processing (N400). Moreover, correlational analysis confirmed that the ERP responses isolated using the two denoising approaches were significantly similar even at the individual level.

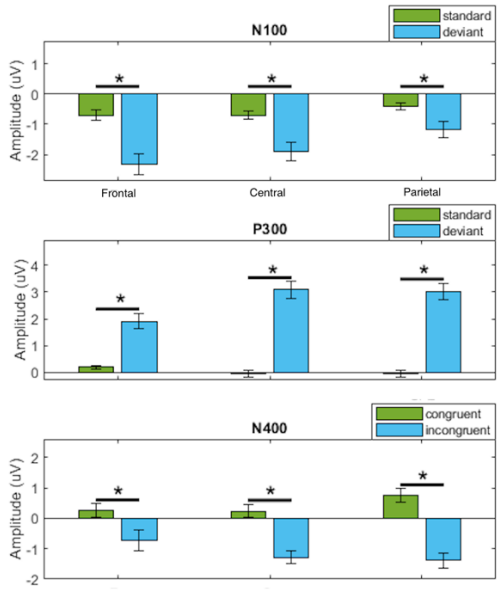


Fig. 4. Comparison of amplitudes across experimental conditions for ERP waveforms derived with EMD-TM technique. ERP waveform amplitudes in time windows of interest across experimental conditions for tone and word stimuli. Results shown as mean ± std. error. *p<0.05.

A major advantage of the EMD-TM technique is the potential ability to denoise individual-channel EEG data, in contrast with many of the existing approaches (e.g. ICA or PCA) that require multiple input sources. With the increasing use of low-density EEG systems for real world applications (e.g. electrodes located in and around the ears [18]), this new approach presents a distinct advantage by not requiring additional electrodes to act as reference measures for the artefacts (e.g. electrodes around the eyes for ocular recording). Furthermore, since EMD-TM works by focusing on the signal of interest (i.e. features of ERP for retention), it is well-suited for applications in realistic settings where the level and types of noise sources can vary quite widely compared to well-controlled

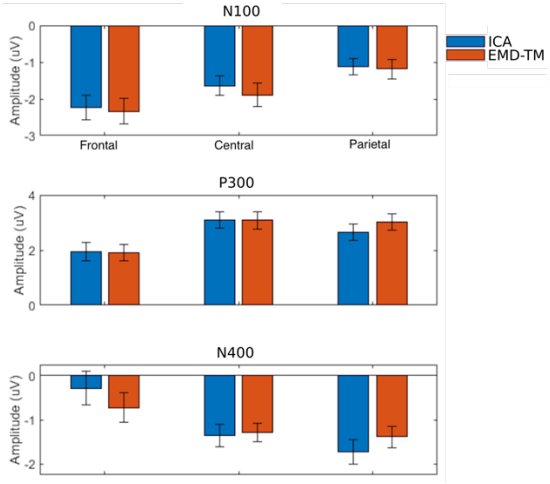


Fig. 5. Mean amplitude of ERP components. Amplitude of response to *deviant* and *incongruent* stimuli generated by ICA (blue) and EMD-TM (orange) denoising approaches. Results shown as mean ± SE. No statistically significant differences were observed between the two approaches.

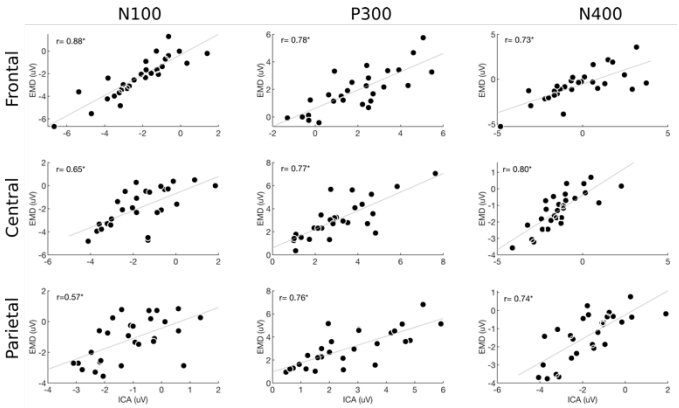


Fig. 6. Correlational analysis of individual-level ERP amplitudes. Mean amplitude of ERP components generated using the ICA and EMD-TM denoising techniques. Correlation coefficient values shown within each plot. *p<0.05.

laboratory environments [14], such as the patient bedside for clinical brain function assessment or the flight cockpit for pilot cognitive state monitoring. The new EMD-TM technique is thus a general purpose denoising approach in that it does not focus on specific types of artefacts/noise to remove in contrast to many other techniques [14], [19], [20].

While the results of the current study are promising, a few caveats need to be highlighted. First, the study used EMD, which can face challenges in high noise environments. Additional investigations are needed to better understand the noise limits at which the current technique may perform sub-optimally, and a simple extension of EMD called ensemble empirical mode decomposition may be used for those high noise situations [11], [21]. Second, in this study, data from the same system was used within a leave-one-out approach for generation of the template and testing the denoising technique. However, future work needs to examine this approach using data from different systems for generation of the template and the test data to be denoised. And lastly, in this study, the results of the EMD-TM technique were compared to the results of the ICA approach. The ICA approach was chosen for comparisons since it represents a technique that is widely applied in this area and requires minimal assumptions and choice of parameters. However, future works should also compare the EMD-TM approach to other denoising techniques besides ICA.

V. Conclusion

This study presents a novel denoising technique that combines EMD with template matching. Results showed that the new EMD-TM technique is capable of operating on individual channel EEG data, and performs comparably to well-established denoising approaches at both the whole-epoch and individual ERP feature levels. With the increasing use of low-density EEG systems for brain function assessments applications in realistic environments, this novel technique provides an avenue to denoise the data and isolate ERP features of interest. As such, this new technique may enable more widespread use of ERP assessments beyond research laboratory environments.

ACKNOWLEDGMENTS

The authors would like to thank the participants for volunteering for this study. The authors would like to acknowledge Mary Carmen Graham and Qun Gao for their administrative and technical support. The authors confirm that this work was undertaken in the absence of any conflict of interest.

REFERENCES

[1] P. L. Nunez and R. Srinivasan, Electric Fields of the Brain: The neurophysics of EEG. 2009.

[2] S. Ghosh Hajra et al., "Developing brain vital signs: Initial framework for monitoring brain function changes over time," Front. Neurosci., vol. 10, 2016.

[3] C. C. Duncan et al., "Event-related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400," Clin. Neurophysiol., vol. 120, no. 11, pp. 1883–1908, Nov. 2009.

[4] L. Sculthorpe-Petley et al., "A rapid event-related potential (ERP) method for point-of-care evaluation of brain function: Development of the Halifax Consciousness Scanner," J. Neurosci. Methods, vol. 245, pp. 64–72, 2015.

[5] R. C. N. D'Arcy, S. Ghosh Hajra, C. Liu, L. D. Sculthorpe, and D. F. Weaver, "Towards Brain First-Aid: A Diagnostic Device for Conscious Awareness," Biomed. Eng. IEEE Trans., vol. 58, no. 3, pp. 750–754, 2011.

[6] S. J. Luck, An introduction to the event-related potential technique. MIT press, 2014.

[7] T.-P. JUNG et al., "Removing electroencephalographic artifacts by blind source separation," Psychophysiology, vol. 37, no. 2, pp. 163–178, 2000.

[8] M. E. Davies and C. J. James, "Source separation using single channel ICA," Signal Processing, vol. 87, no. 8, 2007.

[9] G. Gratton, M. G. . Coles, and E. Donchin, "A new method for off-line removal of ocular artifact," Electroencephalogr. Clin. Neurophysiol., vol. 55, no. 4, pp. 468–484, Apr. 1983.

[10] P. He, G. Wilson, and C. Russell, "Removal of ocular artifacts from electro-encephalogram by adaptive filtering," Med. Biol. Eng. Comput., vol. 42, no. 3, pp. 407–412, 2004.

[11] M. K. Islam, A. Rastegarnia, and Z. Yang, "Methods for artifact detection and removal from scalp EEG: A review," Neurophysiol. Clin. Neurophysiol., vol. 46, no. 4–5, pp. 287–305, 2016.

[12] C. C. Liu, S. Ghosh Hajra, S. Fickling, G. Pawlowski, X. Song, and R. C. N. D'Arcy, "Novel Signal Processing Technique for Capture and Isolation of Blink-Related Oscillations Using a Low-Density Electrode Array for Bedside Evaluation of Consciousness," IEEE Trans. Biomed. Eng., 2019.

[13] N. Williams, S. Nasuto, and J. Saddy, "Evaluation of Empirical Mode Decomposition for Event-Related Potential Analysis," EURASIP J. Adv. Signal Process., 2011.

[14] K. T. Sweeney, S. F. McLoone, and T. E. Ward, "The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique," IEEE Trans. Biomed. Eng., vol. 60, no. 1, 2013.

[15] G. M. Pawlowski et al., "Brain Vital Signs: Expanding From the Auditory to Visual Modality," Front. Neurosci., vol. 12, p. 968, 2019.

[16] C. C. Liu, S. G. Hajra, X. Song, S. M. Doesburg, T. P. L. Cheung, and R. C. N. D'Arcy, "Cognitive loading via mental arithmetic modulates effects of blink-related oscillations on precuneus and ventral attention network regions," Hum. Brain Mapp., vol. 40, no. 0, pp. 377–393.

[17] S. Ghosh Hajra, C. C. Liu, X. Song, S. D. Fickling, T. P. L. Cheung, and R. C. N. D'Arcy, "Accessing knowledge of the 'here and now': a new technique for capturing electromagnetic markers of orientation processing.," J. Neural Eng., vol. 16, no. 1, p. 016008, Feb. 2019.

[18] S. Debener, R. Emkes, M. De Vos, and M. Bleichner, "Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear," Sci. Rep., 2015.

[19] M. Rashed-Al-Mahfuz, M. R. Islam, K. Hirose, and M. K. I. Molla, "Artifact suppression and analysis of brain activities with electroencephalography signals," Neural Regen. Res., vol. 8, no. 16, 2013.

[20] R. Patel, K. Gireesan, S. Sengottuvel, M. P. Janawadkar, and T. S. Radhakrishnan, "Common Methodology for Cardiac and Ocular Artifact Suppression from EEG Recordings by Combining Ensemble Empirical Mode Decomposition with Regression Approach," J. Med. Biol. Eng., vol. 37, no. 2, pp. 201–208, Apr. 2017.

[21] S. D. Fickling, C. C. Liu, R. C. N. D'Arcy, S. Ghosh Hajra, and X. Song, "Good data? the EEG Quality Index for Automated Assessment of Signal Quality," in 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2019, 2019.

A Low SWaP-C Radar Altimeter Transceiver Design for Small Satellites

Ozan Dogan[#], Faruk Uysal[#], Peter Hoogeboom[*], Paco Lopez Dekker[*], Yuanhao Li[*]
[#]TU Delft EWI, Netherlands
[*]TU Delft CITG, Netherlands
{[1]o.dogan@tudelft.nl }

***Abstract* — This paper discusses the design details of a high resolution, low "Size, Weight, Power and Cost" (SWaP-C) radar altimeter (RA) system. Operating frequency of the radar is chosen within the Ka-band to achieve the desired size and weight requirements, that are highly demanded for the small satellite missions in a cost-efficient way. We propose a system design such that, an intended radar altimeter can be built by using the Commercial off the Shelf (COTS) components. The simulation results show that the proposed RA has high potentiality for realization.**

***Keywords* — Radar design, mission concept, radar altimeter**

I. Introduction

A radar altimeter is known to be an elegant tool for estimation of the sea surface height, significant wave height, and wind speed. In the radar system engineering point of view, these parameters are corresponding to the measurement of range, the slope of the return signal and the strength of return signal respectively. Moreover, radar altimeters have great potential in obtaining the reliable knowledge of ocean currents, annual sea level rise, the kinetic energy of ocean circulation and the interaction between the ocean and the ice sheet and glaciers regions, which have been verified by different European Space Agency (ESA) and National Aeronautics and Space Administration (NASA) spaceborne radar altimeters missions (e.g. SWOT, Cryosat-2 [1], Sentinel-3, etc ...).

The traditional altimeters are mostly operating in the Ku [2] and C bands [3]. It is easier to find COTS components operating in these bands. However, the size of the components like antenna and RF filters is much bigger, thus increase also the weight; for i.e., the weight of CryoSAT is 62 kg [4] that is quite a lot to be launched in a small cubesat. Increasing the operating frequency makes the antenna and RF circuitry smaller but unfortunately results in higher cost.

To the best of our knowledge, the highest operating frequency implemented for a radar altimeter is within W-band [5]. However in this case, the prices of the components are too high, and many components especially the power amplifiers are less likely to be found as COTS. Thus custom design of these components is required.

The Ka-band is a good compromise to achieve a low SWaP-C radar altimeter solution from a microwave design point of view. Additionally, this band has benefits to oceanic observations near coastal regions or near the regions between sea ice to inland water. Moreover, Ka-band is less affected by the ionospheric attenuation than the Ku-band.

There is a growing interest in the Ka band radar altimeters. To the best of our knowledge, there are at least three different Ka-band radar altimeter projects. Altika is a Nadir looking altimeter that is launched in 2013 [6] with the France and India collaboration. NASA/CNES wideswath altimetry mission includes a Ka-band Radar Interferometer that is planned to be launched in 2020 [7]. Lastly, ESA has also plans for launching a Ka-band radar altimeter in the next Sentinel 6 mission in 2020 [8]. A comparison of different altimeters are presented in Table 1. The SAR missions are also within the area of interest of the radar altimeter missions [9].

There is an ongoing project at ESA about the investigation of a Cubesat altimeter constellation. A constellation is foreseen to greatly help in understanding the decorrelation times and the dynamic changing of the ocean mesoscale and sub-mesoscale processes, with the system's ultimate goal being a daily worldwide coverage. Instead of a single high power, huge and heavy satellite, low SWaP-C nodes are highly demanded to achieve multiple goals.

In this work, a Ka-band low SWaP-C radar altimeter RF transceiver that is demanded for a cubesat constellation is proposed. In the next section, the system specifications of the proposed radar will be presented. Then the system design will be explained in detail. The following sections discuss the onboard processing, results, and conclusion.

II. System Specifications

The system-level parameters are given in Table 2. There will be two operational modes: Namely, the traditional mode where the pulse repetition frequency (PRF) is 4kHz and the SAR mode in which the PRF is 18.7kHz. It is wise to note that in the radar altimeter jargon "Delay-Doppler" term is used instead of unfocussed SAR [12]. In our design, the peak power is aimed to be 2W in the output of the transmitter. An antenna with a quite narrow beamwidth must be chosen to achieve a high gain (aimed 48dBi) with an antenna width of 1.15m. The bandwidth of the chirp is set to 320MHz which is comparable with the state of the art altimeters. The expected receiver signal power after the antenna is within -127dBm to -87dBm. This is a quite narrow dynamic range that relaxes the analog to digital conversion depth. The expected noise figure (NF) is

978-1-7281-9615-2/20/$31.00

Table 1. Comparison of current altimeters that are installed in the last 30 years.

Name	PRF in Hz and Band	Pulse Width, us	Peak Power, W	Antenna Beamwidth, Degree	Bandwidth
ERS-1, 2	1020 (Ku)	20	50	1.3	330
Topex/Poseidon [10]	4200(Ku), 1220 (C)	102.4	237	1.1	320
Envisat	1795 (Ku), 449 (S)	20	161	1.29 (Ku), 5.5(S)	320
Jason-1, Poseidon 3	2060 (interlaced, Ku and C)	105	7	1.28 (Ku), 3.4 (C)	320
CryoSat-2 [11], SIRAL, Ku	1970 (LRM*), 17.8k (SAR, SARIn)	44.8	25	1.07 * 1.2	350
Sentinel-3	1920, Ku/C	20	-	-	-
SARAL/Altika [6]	3800, Ka	110	2	0.5	500
SWOT,Ka [7]	2 x 4420 Hz	4.5	1500	0.15*2.27	200
Proposed RA	4000, 18.7k (SAR) Ka	48	2	0.5	320

* LRM: Low Resolution Mode

Table 2. System Specifications

Parameter	Value	Unit
Bandwidth	320	MHz
Center Frequency	35.6	GHz
Pulse Width	48	us
PRF	4, 18.7(SAR)	kHz
Output Power	33	dBm
Rx input power	[-127,-87]	dBm
Rx Noise Figure	5	dB
IF Bandwidth after dechirping	7	MHz

less than 5dB. The antenna footprint will be less than 15km which results in a 76 meters scene length in range direction.

The radar altimeter receivers use deramp (stretch) processing to reduce the ADC‘s bandwidth requirements since the duration of the pulse to scan the whole scene is much shorter than the pulse width. Stretch processing generates the beat frequency signal in the receiver which is selected less than 7MHz for the proposed design.

III. RF Transceiver Design

The RF transceiver design is illustrated in Figure 1. The final RA will work as a pulsed radar that includes a circulator and a single antenna. However, to verify and validate the system in the laboratory and on-site where the range is quite short compared to the spaceborne geometry, the presented design includes two antennas instead of a circulator and a single antenna.

The waveform is generated by the Direct Digital Synthesizer (DDS). A two-channel DDS is selected to generate IQ channels without a hybrid splitter. To fulfill the 320MHz system bandwidth requirement, the DDS is chosen to be 500MSPS. The output of DDS is low pass filtered and given as IF input to the IQ (Inphase and Quadrature) mixer. The DDS-Filter-Mixer structure can be alternatively designed by using the DDS as a reference to the Phase Lock Loop (PLL) that drives a Voltage Controlled Oscillator (VCO). So in this case, IQ mixer (MIX IF) and the low pass filter will be no more required. In the final version of the altimeter, this option will be implemented. However, to support also the Arbitrary Waveform Generators (AWG) that can be used for various experiments, the current design will be implemented for the laboratory tests.

The center of the intermediate frequency (IF) is chosen as 2GHz. The selection of IF is a compromise between sideband suppression and the availability of the components. The Local Oscillator (LO) input of the second mixer is 33.6GHz. So a bandpass filter is required to have a center frequency of 35.6 GHz and a stopband at 31.6 GHz; yielding a reasonable Q-factor. The output of the first mixer is given as input to the variable attenuator to fine-tune the transmitted signal power in order not to saturate the successive amplifier. The IF amplifier is used to provide sufficient power to the local oscillator (LO) of the receiver mixer (MIX IF Rx).

The selected RF mixer (MIX RF) is IQ. A radar altimeter is not required to have an IQ mixer as the beat frequency is always positive. An important disadvantage of the IQ mixer is the need for higher power in the LO input comparing with the LO requirement of a single channel one. For the Ka band mixer in this design, the required LO power is 18dBm, which is quite high. However, to achieve a high-quality sideband suppression capability, the mixer is chosen to be IQ. To generate this power, at least one additional amplifier (RA LO2) is inserted as it is seen at the bottom of Figure 1.

The variable attenuator (Var ATTN RF) at the output of the RF mixer is used to guarantee the succeeding amplifiers not to be saturated and also to compensate the power. The technology for driver amplifier and the power amplifier are chosen as solid-state. The power amplifier is selected to have a P1dB of higher than 33dBm to achieve the required output power. The bandpass filter, (BPF RF) is used for suppression of the sidebands of the mixer and harmonics of the amplifiers.

The transmitted signal is coupled to the receive path for the calibration of the system. This calibration aims to extract the impulse response of the transceiver. In this path, the transmitted signal is delayed and attenuated to inject a simulated target to the receive path. There is also a power detector in the calibration path to control the output power and fine-tune the variable attenuator.

The receiver chain includes a limiter, BPF, LNA and a second BPF. By doubling the filters, the NF can be decreased more with the cost of a bigger size of the PCB. The mixer (MIX RF) is the same as the one in the transmitter. The output signal is combined with a hybrid combiner (HYB COMB) and given as input to the amplifier. This amplifier before the mixer has also a positive impact on the final NF. The remaining

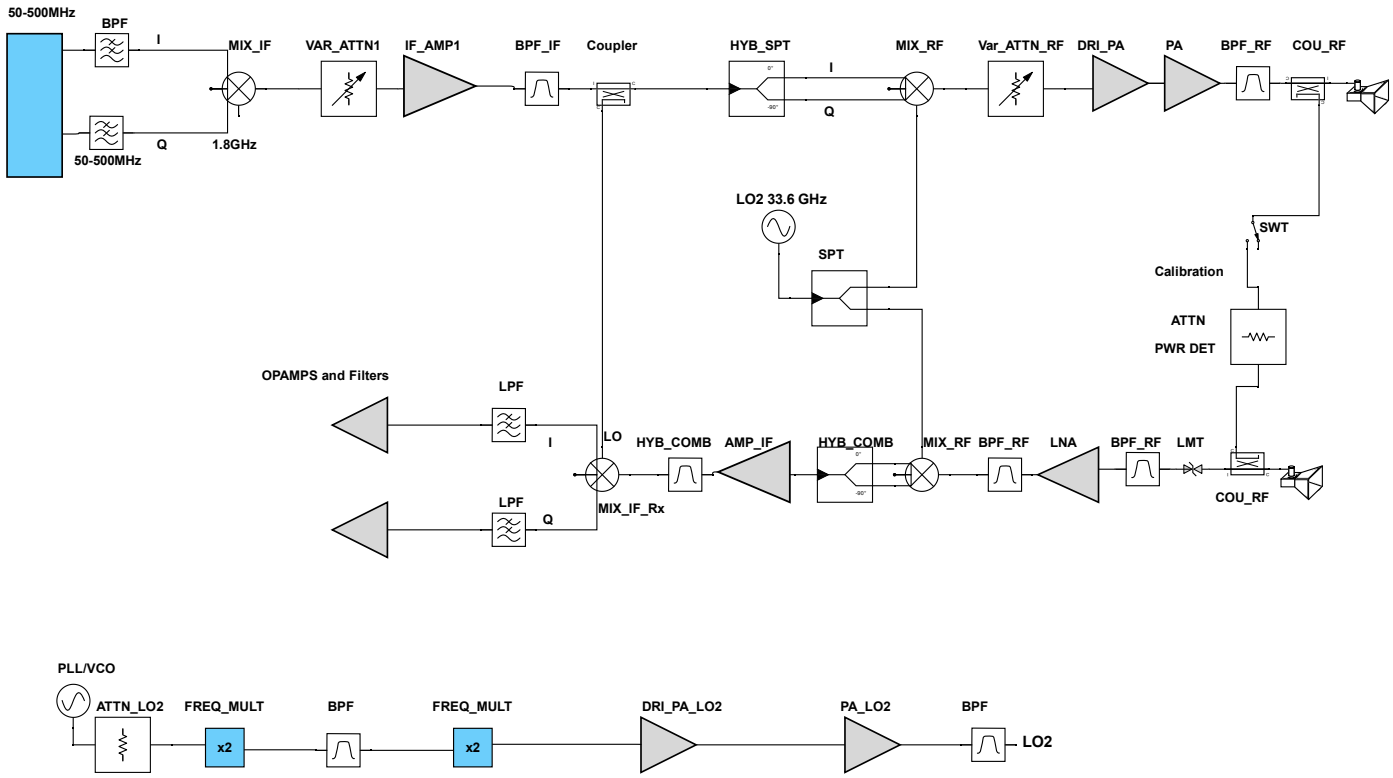


Fig. 1. Radar Altimeter Transceiver Design

part of the receiver is conversion to baseband of the I and Q channels, low pass filtering and amplification of the signal to achieve the required ADC dynamic range.

The expected size of the PCB for the RF transceiver and the signal processing board will be less than 4 by 10cm*10cm. The PCB's can be realized as multilayer that results in shorter heights. The weight of the PCB and the electromagnetic shields will be less than 2 kg. The power required during transmission for the PA is 12W, for the other RF amplifiers is 6W where the duty cycle is 5%. The power required for the whole radar is expected to be less than 30 W. The cost of the PCBs and the all components are estimated to be less than 20k Euro. Hence, the proposed design can be defined as low SWaP-C.

The space qualification of the RF transceivers is under discussion. To the best of our knowledge, there is no consensus on if COTS products are needed to be selected as space-qualified to develop a space-qualified product. In the proposed system, the RF amplifiers (DRI PA, PA, LNA) are chosen as space-qualified. Besides, most of the COTS products have specifications that fit the extreme conditions of the space.

IV. On Board Processing

The onboard processing aims to control the transceiver, to digitize IQ signal and to achieve the DSP functionality. The control of the transceiver includes the fine-tuning of the waveform duration to achieve the beat frequency within the bandwidth. For this purpose, the received signal is firstly digitized. The required minimum number of samples to cover the 150m range is 320 that is between 256 and 512. A 512 sample complex FFT at the PRF is required to be performed and the squares of I and Q must be summed up to compute the signal amplitude. The input of the tracking algorithm is the detection results over the averaged signal amplitude and the Digital Elevation Model of the scene. The tracked range is given as input for chirp duration estimation so that the height measurement will be in the middle of the range window.

The onboard processing will also include the control mechanisms for the states of the system. The calibration will be the initial state of the system. In this state the transmitted power will be detected, the PLL lock signals will be checked and the system impulse response will be measured. The next stage is the search state in which, the height will be measured with a narrow band chirp. Once the height is estimated roughly, the band will be increased in a loop like focusing a telescope in a stepwise manner. The height accuracy can be increased by interpolation techniques [13]. Once the highest bandwidth is achieved, the system will be in the lock state. If there is an inconsistency of the system, the lock state will convert back to the search state. In order to achieve the onboard functionality, space proven FPGAs will be used.

V. Simulation Results

In this section, the simulation results of the RA system will be discussed. The system is simulated by using the AWR Microwave circuit design tool. In Figure 2 the transmitter chain signal power analysis is presented. The brown line that has the lowest values represents the power of the transmitted signal, the pink line is P1dB and the blue line is the IP3. As it is seen

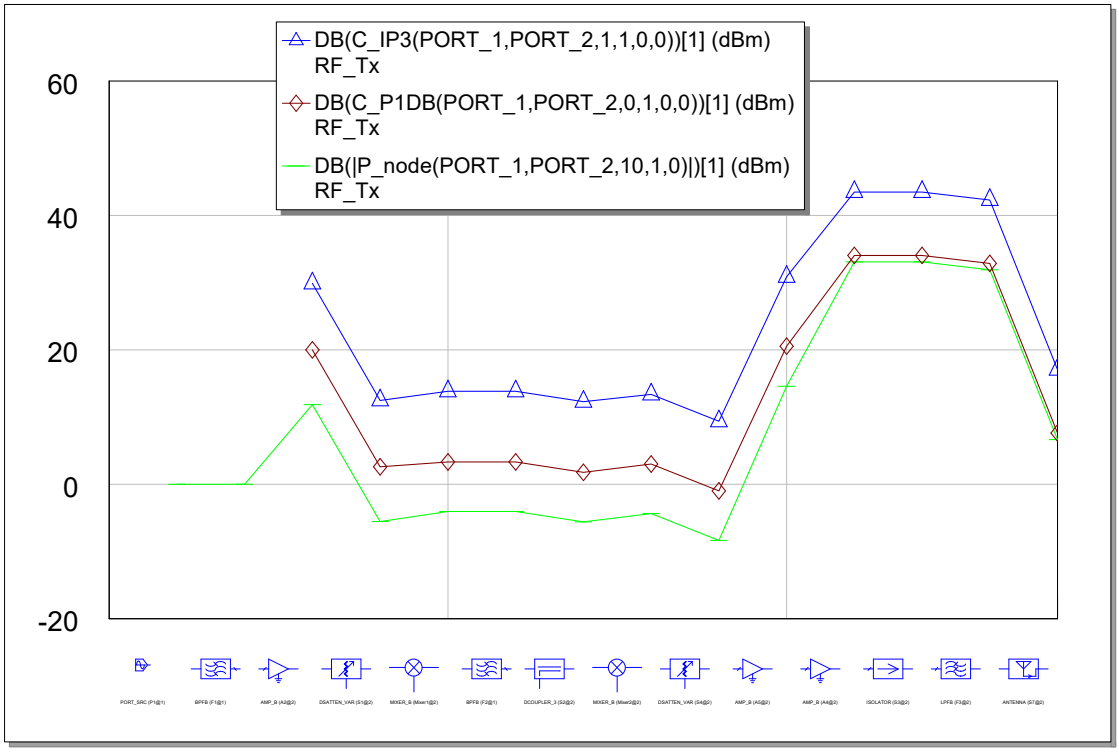


Fig. 2. Transmitter chain RF budget analysis. (dBm)

the transmitted signal is always lower than the others and is the same with P1dB only at the output of the transceiver. The output power is 34.4dBm that is 1.4dB higher than the system requirement.

In Figure 3 the NF of the receiver chain is presented. The NF requirement was 5dB, and it is seen that 1dB is saved. The bandpass filter after the LNA and the amplifier after the mixer are inserted to the chain to keep the NF in these lower values. This is a compromise between the size, power and NF.

VI. Conclusion

This paper presents a Ka-band RA transceiver design and discusses the design choices in each stage. All the components are selected as COTS, which can be easily found in the market, to achieve not only a low SWaP but also a low-cost solution. The simulation results show that the design fulfills the system-level requirements. The next step is to implement and validate the design by laboratory and cite tests.

Acknowledgment

The work has been carried out within the framework of the NL-RIA "Dutch Knowledge Network on Radar Instruments and Applications" project funded by the Netherlands Organisation for Scientific Research (NWO). We also thank Reindert Grooters from TNO for his very useful contributions to the RF design.

References

[1] A. S. R. L. Tilling, A. Ridout, "Estimating arctic sea ice thickness and volume using cryosat-2 radar altimeter data," *Advances in Space Research*, vol. 62, no. 6, pp. 1203–1225, 2018.

[2] H. T. W. Kong, J. Chong, "Performance analysis of ocean surface topography altimetry by ku-band near-nadir interferometric sar," *Remote Sens.*, 2017.

[3] K. Xu, H. Liu, P. Liu, and X. Yu, "System techniques and capabilities of the hy-2a altimeter," in *2014 IEEE Geoscience and Remote Sensing Symposium*, July 2014, pp. 5171–5174.

[4] ESA, "Cryosat mission and data description," *ESA Report, Doc. No.: CS-RP-ESA-SY-0059*, 2007.

[5] L. Li, G. M. Heymsfield, L. Tian, and P. E. Racette, "Measurements of ocean surface backscattering using an airborne 94-GHz cloud radar—implication for calibration of airborne and spaceborne W-band radars," *Journal of Atmospheric and Oceanic Technology*, vol. 22, no. 7, pp. 1033–1045, 2005. [Online]. Available: https://doi.org/10.1175/JTECH1722.1

[6] J. Verron, P. Sengenes, and J. Lambin, "The SARAL/AltiKa satellite altimetry mission," *Journal Marine Geodesy*, Volume 38,2015.

[7] R. Fjørtoft, J. Gaudin, N. Pourthié, J. Lalaurie, A. Mallet, J. Nouvel, J. Martinot-Lagarde, H. Oriot, P. Borderies, C. Ruiz, and S. Daniel, "KaRIn on SWOT: Characteristics of Near-Nadir Ka-Band Interferometric SAR Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 4, pp. 2172–2185, April 2014.

[8] F. Ardhuin, P. Brandt, and L. Gaultier, "Kim, a candidate satellite mission exploring global ocean currents and waves," *Frontiers in Marine Science*, April 2019.

[9] F. X. Zhanyu Zhu, Hai Zhang, "Raw signal simulation of synthetic

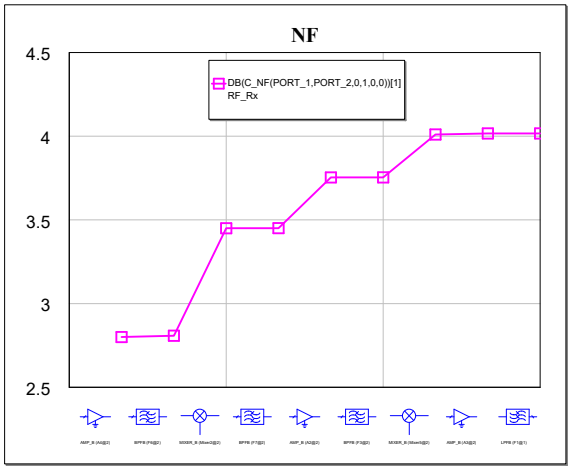


Fig. 3. Receiver chain Noise figure. (dB)

aperture radar altimeter over complex terrain surfaces," *Radio Science*, vol. 55, no. 2, 2020.

[10] G. Carayon, N. Steunou, J. Courrière, and P. Thibaut, "Poseidon-2 radar altimeter design and results of in-flight performances special issue: Jason-1 calibration/validation," *Marine Geodesy*, vol. 26:3-4, pp. 159–165, Nov. 2003.

[11] D. Wingham, C. Francis, and S. Baker, "Cryosat: A mission to determine the fluctuations in earth's land and marine ice fields," *Elsevier Advances in Space Research*, 2006.

[12] R. K. Raney, "Advances in oceanic radar altimetry:cryosat and beyond," *ESA Workshop*, 2014.

[13] O. Dogan and I. Erer, "High resolution radar imaging using gpof based data extrapolation," *IEEE International Geoscience and Remote Sensing Symposium, Alaska Anchorage*, 2004.

An Automated System to Limit COVID-19 Using Facial Mask Detection in Smart City Network

Mohammad Marufur Rahman
Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
inkmarufnayem@gmail.com

Md. Motaleb Hossen Manik
Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
mkmanik557@gmail.com

Md. Milon Islam
Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
milonislam@cse.kuet.ac.bd

Saifuddin Mahmud
Advanced Telerobotics Research Lab
Computer Science
Kent State University
Kent, Ohio, USA
smahmud2@kent.edu

Jong-Hoon Kim
Advanced Telerobotics Research Lab
Computer Science
Kent State University
Kent, Ohio, USA
jkim72@kent.edu

***Abstract*— COVID-19 pandemic caused by novel coronavirus is continuously spreading until now all over the world. The impact of COVID-19 has been fallen on almost all sectors of development. The healthcare system is going through a crisis. Many precautionary measures have been taken to reduce the spread of this disease where wearing a mask is one of them. In this paper, we propose a system that restrict the growth of COVID-19 by finding out people who are not wearing any facial mask in a smart city network where all the public places are monitored with Closed-Circuit Television (CCTV) cameras. While a person without a mask is detected, the corresponding authority is informed through the city network. A deep learning architecture is trained on a dataset that consists of images of people with and without masks collected from various sources. The trained architecture achieved 98.7% accuracy on distinguishing people with and without a facial mask for previously unseen test data. It is hoped that our study would be a useful tool to reduce the spread of this communicable disease for many countries in the world.**

Keywords—Facial Mask Detection, COVID-19, Deep Learning, Convolutional Neural Network, Smart City.

I. INTRODUCTION

A new strain which has not previously been identified in humans is novel coronavirus (nCoV). Coronaviruses (CoV) are a wide group of viruses which cause illness that range from colds to deadly infections like Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS) [1]. The first infected patient of coronavirus has been found in December 2019. From that period, COVID-19 has become a pandemic all over the world [2]. People all over the world are facing challenging situations due to this pandemic. Every day a large number of people are being infected and died. At the time of writing this paper, almost 16,207,130 infected cases have been confirmed where 648,513 are death [3]. This number is increasing day by day. Fever, dry cough, tiredness, diarrhea, loss of taste, and smell are the major symptoms of coronavirus which is declared by the World Health Organization (WHO) [4]. Many precautionary measures have been taken to fight against coronavirus. Among them cleaning hands, maintaining a safe distance, wearing a mask, refraining from touching eyes, nose, and mouth are the main, where wearing a mask is the simplest one. COVID-19 is a disease that spread from human to human which can be controlled by ensuring proper use of a facial mask. The spread of COVID-19 can be limited if people strictly maintain social distancing and use a facial mask. Very sadly, people are not obeying these rules properly which is speeding the spread of this virus. Detecting the people not obeying the rules and informing the corresponding authorities can be a solution in reducing the spread of coronavirus.

A face mask detection is a technique to find out whether someone is wearing a mask or not. It is similar to detect any object from a scene. Many systems have been introduced for object detection. Deep learning techniques are highly used in medical applications [5], [6]. Recently, deep learning architectures [7] have shown a remarkable role in object detection. These architectures can be incorporated in detecting the mask on a face. Moreover, a smart city [8] means an urban area that consists of many IoT sensors to collect data. These collected data are then used to perform different operations across the city. This includes monitoring traffic, utilities, water supply network, and many more. Recently, the growth of COVID-19 can be reduced by detecting the facial mask in a smart city network.

This paper aims at designing a system to find out whether a person is using a mask or not and informing the corresponding authority in a smart city network. Firstly, CCTV cameras are used to capture real-time video footage of different public places in the city. From that video footage, facial images are extracted and these images are used to identify the mask on the face. The learning algorithm Convolutional Neural Network (CNN) is used for feature extraction from the images then these features are learned by multiple hidden layers. Whenever the architecture identifies people without face mask this information is transferred through the city network to the corresponding authority to take necessary actions. The proposed system appraised promising output on data collected from different sources. We also represented a system that can ensure proper enforcement of the law on people who are not following basic health guidelines in this pandemic situation.

The remainder of the paper is arranged accordingly. The most recent works for facial mask detection is described in

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Section II. In Section III, the proposed methodology for developing the whole system is described. Section IV analyses the results obtained from the developed system. The conclusion is drawn in Section V. Lastly, the limitations with potential further works are depicted in Section VI.

II. Related Works

In the meantime, many systems have been developed for COVID-19 in smart city networks. BlueDot and HealthMap services have been introduced in [9]. BlueDot method was first used to mark the cluster of unusual pneumonia in Wuhan which finally detected the disease as a pandemic. It also predicted that the virus would spread from Wuhan to Bangkok, Taipei, Singapore, Tokyo and Hong Kong. HealthMap service, based on San Francisco, spotted the patients with a cough which is the initial sign of COVID-19, using Artificial Intelligence (AI) and big data. A study on using facemask to restrict the growth of COVID-19 is introduced in [10]. The study indicated that the masks that are adequately fit, effectively interrupt the spread of droplets expelled when coughing or sneezing. Masks that are not perfectly fitted, also capable of retaining airborne particles and viruses. Allam and Jones [11] proposed a framework on smart city networks focusing on how data sharing should be performed during the outbreak of COVID-19. The proposed system discussed the prospects of Urban Health Data regarding the safety issues of the economy and national security. In the system, the data is collected from various points of the city using sensors, trackers, and from laboratories.

A face mask detecting model named RetinaFaceMask combining with a cross-class object removal algorithm is proposed by Jiang et al. [12]. The developed model includes one stage detector consisting feature pyramid network that results in slightly higher precision and recall than the baseline result. For reducing the shortage of datasets, they have applied transfer learning, a well-known deep learning technique. Gupta et al. [13] proposed a model to enforce the social distance using smart city and Intelligent Transportation System (ITS) during COVID-19 pandemic. Their model described the deploying sensors in different places of the city to monitor the real-time movement of objects and offered a data-sharing platform. A noticeable contribution of a smart city in controlling the spread of coronavirus in South Korea is explained by Won Sonn and Lee [14]. A time-space cartographer speeded up the contact tracking in the city including patient movement, purchase history, cell phone usages, and cell phone location. Real-time monitoring has been carried out on CCTV cameras in the hallways of residential buildings.

Singh et al. [15] put their focus on how IoT can fight against COVID-19. The developed system emphasizes on inter-connected devices or operations to track the patients along with wary cases. A well-informed group using inter-connected devices is formed to identify the clusters significantly. A remarkable pandemic control model without lockdown in a smart city has been outlined by Sonn et al. [16]. The patients have been interviewed and their past movement has been monitored. They have claimed that some patients tried to conceal about their past mobility but real-time tracking system found the exact information. Jaiswal et al. [17] proposed a way to minimize the risk during COVID-19. Their proposed model used the position of technology to track infected people. Drones and Robot technologies have been applied as medical personnel for providing adequate services to infected people. The development of smart cities under COVID-19 and controlling the pandemic in China has been reviewed by Wang et al. [18]. The continuous supply of essential materials and contactless logistic distribution of systems to society made the way to reduce the spread of coronavirus. ITS and real-time map reflection methods have been used to block the movement of vehicles during the pandemic. In addition, driverless vehicles have been used to monitor the scenarios across the city.

III. Methodology

We proposed an automated smart framework for screening persons who are not using a face mask in this paper. In the smart city, all public places are monitored by CCTV cameras. The cameras are used to capture images from public places; then these images are feed into a system that identifies if any person without face mask appears in the image. If any person without a face mask is detected then this information is sent to the proper authority to take necessary actions. The block diagram of the developed framework is depicted in Fig. 1. All the blocks of the developed system are described as follows.

A. Image Preprocessing

The images captured by the CCTV cameras required preprocessing before going to the next step. In the preprocessing step, the image is transformed into a grayscale image because the RGB color image contains so much redundant information that is not necessary for face mask detection. RGB color image stored 24 bit for each pixel of the image. On the other hand, the grayscale image stored 8 bit for each pixel and it contained sufficient information for classification. Then, we reshaped the images into (64×64) shape to maintain uniformity of the input images to the architecture. Then, the images are normalized and after normalization, the value of a pixel resides in the range from 0 to 1. Normalization helped the learning algorithm to learn faster and captured necessary features from the images.

B. Deep Learning Architecture

The deep learning architecture learns various important nonlinear features from the given samples. Then, this learned architecture is used to predict previously unseen samples. To train our deep learning architecture, we collected images from different sources. The architecture of the learning technique highly depends on CNN. All the aspects of deep learning architecture are described below.

i) Dataset Collection: Data from two different sources [19], [20] are collected for training and testing the model. We collected a total of 858 images of people with masks and 681 images of people without a mask. For training purposes, 80% images of each class are used and the rest of the images are utilized for testing purposes. Fig. 2 shows some of the images of two different classes.

ii) Architecture Development: The learning model is based on CNN which is very useful for pattern recognition from images [21]. The network comprises an input layer, several hidden layers and an output layer. The hidden layers consist of multiple convolution layers that learn suitable filters for important feature extraction from the given samples. The features extracted by CNN are used by multiple dense neural networks for classification purposes. The architecture of the developed network is illustrated in Table I. The architecture contains three pairs of convolution layers each followed by

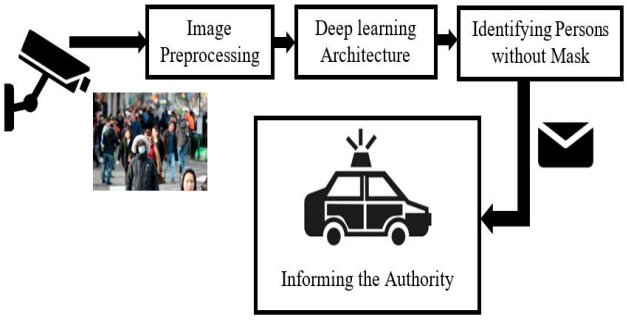


Fig. 1. Block diagram of the proposed system.

People without mask People with mask

Fig. 2. Sample images from the used dataset.

TABLE I. THE ARCHITECTURE OF THE DEEP LEARNING NETWORK

Layer	Type	Kernel	Kernel Size	Output Size
1	Convolution2D	32	(3×3)	(62×62×32)
2	Convolution2D	32	(3×3)	(60×60×32)
3	MaxPooling2D	-	(2×2)	(30×30×32)
4	Convolution2D	32	(3×3)	(28×28×32)
5	Convolution2D	32	(3×3)	(26×26×32)
6	MaxPooling2D	-	(2×2)	(13×13×32)
7	Convolution2D	32	(3×3)	(11×11×32)
8	Convolution2D	32	(3×3)	(9×9×32)
9	MaxPooling2D	-	(2×2)	(4×4×32)
10	Flatten	-	-	512
11	Dense	-	-	100
12	Dropout	-	-	100
13	Dense	-	-	30
14	Dropout	-	-	30
15	Dense	-	-	10
16	Dropout	-	-	10
17	Dense	-	-	2

one max pooling layer. This layer decreases the spatial size of the representation and thereby reduces the number of parameters. As a result, the computation is simplified for the network. Then, a flatten layer reshapes the information into a vector to feed into the dense network. Three pairs of dense and dropout layers learn parameters for classification. The dense layer comprises a series of neurons each of them learn nonlinear features. The dropout layer prevents the network from overfitting by dropping out units. Finally, a dense layer containing two neurons distinguishes the classes.

iii) Screening and Informing the Authority: The main goal of our proposed system is screening persons who are not following guidelines of using a facial mask. The learning architecture identifies whether any input image contains persons without a face mask. If such a person is detected, then this information is sent to the proper authority. The GPS location of the CCTV camera captured the person without a mask along with the image and the exact time is sent via SMS to the corresponding authority. They would come to the locality where the person without a face mask was detected and took necessary actions. If proper actions are taken, then people might not come in public places without a facial mask that would help greatly to limit the growth of COVID-19.

IV. RESULT ANALYSIS

By preserving a reasonable proportion of different classes, the dataset is partitioned into training and testing set. The dataset comprises of 1539 samples in total where 80% is used in training phase and 20% is used in testing phase. The training and testing dataset contains 1231 and 308 images respectively. The developed architecture is trained for 100 epochs since further training results cause overfitting on the training data. Overfitting occurs when a model learns the unwanted patterns of the training samples. Hence, training accuracy increases but test accuracy decreases. Fig. 3 and Fig. 4 show the graphical view of accuracy and loss respectively. The trained model showed 98.7% accuracy and AUC of 0.985 on the unseen test data.

In Fig. 3, the accuracy curve of training and testing is shown for about 100 epochs. From Fig. 3, it is realized that the training and testing accuracy are almost identical. This means the model has a decent generalization ability for previously unseen data and it does not cause overfitting of the training data. In Fig. 4, loss curves of training and testing phases are shown. Here, it is evident that the training loss is decreasing over increasing epochs. The testing loss is lower than training loss for about 30 epochs but after that, it started increasing w means the confidence of prediction started decreasing. The testing loss fluctuates between an acceptable range and it falls about at 98th epoch.

Table II represents the confusion matrix of the testing phase. The developed architecture misclassifies only 04 samples out of 308 samples. It classifies 01 sample as with mask while it is in without mask class and classifies 03 samples as without mask while these were in with mask class. The main aim of the system is to identify samples within without mask class and this architecture misclassified only 01 sample of this class that shows the reliability of the developed system.

Fig. 5 depicts the receiver operating characteristic (ROC) curve of the proposed framework. This illustrates the prediction ability of the classifier at different thresholds. Two parameters are plotted in the ROC curve; one is the true positive rate (TPR) and other is the false positive rate (FPR) measured using (1) and (2) respectively. TPR and FPR are calculated for different threshold and these values are plotted as ROC curve. The area under the ROC curve (AUC) measures the performance of the binary classifier for all possible thresholds. The value of AUC ranges from 0 to 1. When a model predicts 100% correct its AUC is 1 and when it predicts 100% wrong then its AUC is 0. The AUC achieved form our classifier is 0.985 that points towards a decent classifier.

$$True\ Positive\ Rate = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (1)$$

$$False\ Positive\ Rate = \frac{False\ Positive}{True\ Negative + False\ Positive} \quad (2)$$

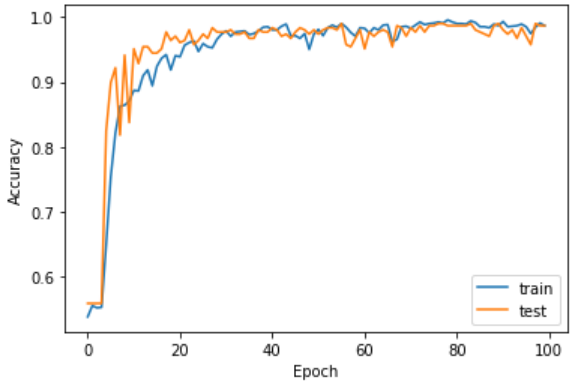


Fig. 3. Accuracy of the developed system for training and testing phase.

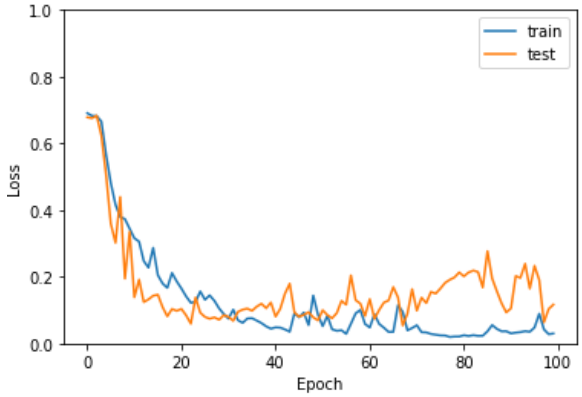


Fig. 4. Loss of the developed system for training and testing phase.

TABLE II. THE CONFUSION MATRIX OF THE DEVELOPED SYSTEM

		Predicted Class	
		Without Mask	With Mask
True Class	Without Mask	134	1
	With Mask	3	170

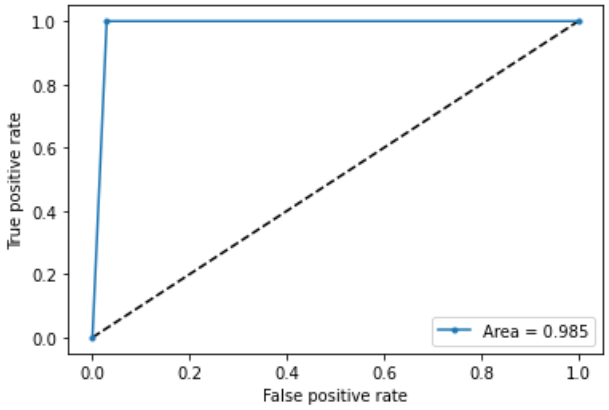


Fig. 5. ROC of the classification network.

V. CONCLUSION

This paper presents a system for a smart city to reduce the spread of coronavirus by informing the authority about the person who is not wearing a facial mask that is a precautionary measure of COVID-19. The motive of the work comes from the people disobeying the rules that are mandatory to stop the spread of coronavirus. The system contains a face mask detection architecture where a deep learning algorithm is used to detect the mask on the face. To train the model, labeled image data are used where the images were facial images with masks and without a mask. The proposed system detects a face mask with an accuracy of 98.7%. The decision of the classification network is transferred to the corresponding authority. The system proposed in this study will act as a valuable tool to strictly impose the use of a facial mask in public places for all people.

VI. LIMITATIONS AND FUTURE WORKS

The developed system faces difficulties in classifying faces covered by hands since it almost looks like the person wearing a mask. While any person without a face mask is traveling on any vehicle, the system cannot locate that person correctly. For a very densely populated area, distinguishing the face of each person is very difficult. For this type of scenario, identifying people without face mask would be very difficult for our proposed system. In order to get the best result out of this system, the city must have a large number of CCTV cameras to monitor the whole city as well as dedicated manpower to enforce proper laws on the violators. Since the information about the violator is sent via SMS, the system fails when there is a problem in the network.

The proposed system mainly detects the face mask and informs the corresponding authority with the location of a person not wearing a mask. Based on this, the authority has to send their personnel to find out the person and take necessary actions. But this manual scenario can be automated by using drones and robot technology [22], [23] to take action instantly. Furthermore, people near to the person not wearing a mask may be alerted by an alarm signal on that location, and displaying the violators face in a LED screen to maintain a safe distance from the person would be a further study.

REFERENCES

[1] WHO EMRO | About COVID-19 | COVID-19 | Health topics. [Online]. Available: *http://www.emro.who.int/health-topics/corona-virus/about-covid-19.html*, accessed on: Jul. 26, 2020.

[2] H. Lau *et al.*, "Internationally lost COVID-19 cases," *J. Microbiol. Immunol. Infect.*, vol. 53, no. 3, pp. 454–458, 2020.

[3] Worldometer, "Coronavirus Cases,". [Online]. Available: *https://www.worldometers.info/coronavirus*, accessed on: Jul. 26, 2020.

[4] L. Li *et al.*, "COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis," *J. Med. Virol.*, vol. 92, no. 6, pp. 577–583, Jun. 2020.

[5] M. Z. Islam, M. M. Islam, and A. Asraf, "A Combined Deep CNN-LSTM Network for the Detection of Novel Coronavirus (COVID-19) Using X-ray Images," *Informatics in Medicine Unlocked*, vol. 20, pp. 100412, Aug. 2020.

[6] L. J. Muhammad, M. M. Islam, S. S. Usman, and S. I. Ayon, "Predictive Data Mining Models for Novel Coronavirus (COVID-19) Infected Patients' Recovery," *SN Comput. Sci.*, vol. 1, no. 4, p. 206, Jun. 2020.

[7] L. Liu *et al.*, "Deep Learning for Generic Object Detection: A Survey," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 261–318, Sep. 2018.

[8] L. Calavia, C. Baladrón, J. M. Aguiar, B. Carro, and A. Sánchez-Esguevillas, "A Semantic Autonomous Video Surveillance System for Dense Camera Networks in Smart Cities," *Sensors*, vol. 12, no. 8, pp. 10407–10429, Aug. 2012.

[9] G. Halegoua, "Smart City Technologies," *Smart Cities*, 2020, doi: 10.7551/mitpress/11426.003.0005.

[10] L. P. Garcia, "Uso de máscara facial para limitar a transmissão da COVID-19," *Epidemiol. e Serv. saude Rev. do Sist. Unico Saude do Bras.*, vol. 29, no. 2, p. e2020023, 2020.

[11] Z. Allam and D. S. Jones, "On the Coronavirus (COVID-19) Outbreak and the Smart City Network: Universal Data Sharing Standards Coupled with Artificial Intelligence (AI) to Benefit Urban Health Monitoring and Management," *Healthcare*, vol. 8, no. 1, p. 46, 2020.

[12] M. Jiang, X. Fan, and H. Yan, "RetinaMask: A Face Mask detector," 2020. [Online]. Available: *http://arxiv.org/abs/2005.03950*.

[13] M. Gupta, M. Abdelsalam, and S. Mittal, "Enabling and Enforcing Social Distancing Measures using Smart City and ITS Infrastructures: A COVID-19 Use Case," 2020. [Online]. Available: *https://arxiv.org/abs/2004.09246.*

[14] J. Won Sonn and J. K. Lee, "The smart city as time-space cartographer in COVID-19 control: the South Korean strategy and democratic control of surveillance technology," *Eurasian Geogr. Econ.*, pp. 1–11, May. 2020.

[15] R. P. Singh, M. Javaid, A. Haleem, and R. Suman, "Internet of things (IoT) applications to fight against COVID-19 pandemic," *Diabetes Metab. Syndr. Clin. Res. Rev.*, vol. 14, no. 4, pp. 521–524, Jul. 2020.

[16] J. W. Sonn, M. Kang, and Y. Choi, "Smart city technologies for pandemic control without lockdown," *Int. J. Urban Sci.*, vol. 24, no. 2, pp. 149–151, 2020.

[17] R. Jaiswal, A. Agarwal, and R. NEGI, "Smart Solution for Reducing the COVID-19 Risk using Smart City Technology," *IET Smart Cities*, vol. 2, pp. 82–88, 2020.

[18] X. Wang, X. Le, and Q. Lu, "Analysis of China's Smart City Upgrade and Smart Logistics Development under the COVID-19 Epidemic," *J. Phys. Conf. Ser.*, vol. 1570, p. 012066, 2020.

[19] Face Mask Detection | Kaggle. [Online]. Available: *https://www.kaggle.com/andrewmvd/face-mask-detection*, accessed on: Jul. 27, 2020.

[20] GitHub-prajnasb/observations. [Online]. Available: *https://github.com/prajnasb/observations*, accessed on: Jul. 27, 2020.

[21] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A Survey of the Recent Architectures of Deep Convolutional Neural Networks," *Artif. Intell. Rev.*, Jan. 2019.

[22] I. S. Cardenas *et al.*, "Telesuit: design and implementation of an immersive user-centric telepresence control suit," in *Proceedings of the 23rd International Symposium on Wearable Computers - ISWC '19*, New York, NY, USA, 2019, pp. 261–266.

[23] D. Y. Kim, I. S. Cardenas, and J.-H. Kim, "Engage/Disengage: Control Triggers for Immersive Telepresence Robots," in *Proceedings of the 5th International Conference on Human Agent Interaction*, New York, NY, USA, 2017, pp. 495–499.

A Light-Based Quantum-Computational Model of Genetics

Dr Pravir Malik
Deep Order Technologies, USA
pravir.malik@deepordertechnologies.com

Abstract— Light imagined to exist at different constant speeds far greater than the known speed of c – 186,000 miles per second - offers a unique view of quanta and quantum computation. Analyses of such light further provides a unique point of view into the origin and possibilities of genetics. Genetics can be seen as having a diverse light-based functional, as opposed to a solely form-based foundation. Genetics can also be perceived as being the output of a persistent quantum-level computation. Such a modeling provides useful hypotheses into the structure of DNA, into processes of constructive and destructive mutation, and heredity. Further the relationship and possible impacts of the quantum-based processes of entanglement and superposition on genetics, and future possibilities due to practically infinite amount of information in antecedent layers of light can be constructed. Finally, such a view may suggest a non-invasive and constructive way in which to influence processes that underly genetics.

Keywords— ***Genetics, Mutation, Quantum Computation, Heredity, Symmetries in Light, Superposition, Entanglement, Genetic-Type Information, Quanta***

I. Introduction

The Cosmology of Light computational model (Malik 2018a, b; 2019, 2020) explored a mathematical structure of Light composed of light traveling at different constant speeds beyond the known speed of light, of c – 186,000 miles per second in vacuum - in the physical universe. Light at different speeds engenders different realities, which will be summarized in Sections II and III, and as a result engenders different kinds of information. This vast variety of information is proposed as being the origin of genetic-type information and genetics.

Light at different speeds, from an imagined speed of infinite miles per second, down to c - 186,000 miles per second in vacuum - creates different though mathematically symmetrical realities based on four underlying properties suggested to be implicit in light. This gives rise to a quaternary-based multi-layered mathematical structure in which the unity of light traveling infinitely fast separates in an increasing display of distinct functionality to create the reality of infinite functional diversity we are familiar with in this reality where light travels at c. This happens through a process of precipitation. This process of precipitation can also be perceived as a "subtle" backbone or strand not unlike the backbone or downward-strand of the double helix DNA structure that animates every living cell. In other words, the very basis of genetic structure may potentially be tied to the process of precipitation of light.

Further, as light slows down the codification implicit in the state where it is imagined travelling infinitely fast begins to further materialize and generates "libraries" of information as it were, that are accessible to all subsequent layers created by "slower" light. These libraries are the bases of genetic information and heredity.

The process of mutation will be seen to vary depending on which layer of light is involved. Massively significant mutations such as took place on the SRY gene and led to the evolution of human from the chimpanzee species (Ridley, 1999) are suggested to be due to layers of light traveling faster than c. By contrast degenerative mutation leading to dysfunction and disease are hypothesized to be linked to the layer of reality so set up by light projected at slower than c speeds or zero speed.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Section IV will review the afore-mentioned genetic concepts from the point of view of a light-based quantum-computational model.

Section V will summarize and conclude with some thoughts on future directions of genetic research in such a light-based quantum-computation model.

II. Light And The Origin Of Genetics

Practically, the finite speed of light implies that light will take a finite amount of time to travel from one point to another. This is significant even when viewed at the atomic scale. If there is a source of electromagnetic radiation in the nucleus or due to the electrons changing orbits around a nucleus that radiation will be experienced only a finite time later. It can be inferred that this phenomenon is related to quanta. Energy of quanta is specified by (1), where 'E' is energy of quanta, 'h' is Planck's constant, $'\upsilon'$ is frequency of the radiation. Hence (1):

$$E = h\upsilon$$

Eq. 1: Energy of Quanta

Further, the speed of light 'c', frequency of radiation $'\upsilon'$, and wavelength of radiation $'\lambda'$, are related by (2):

$$c = \upsilon\lambda$$

Eq. 2: Speed of Light

Combining (1) and (2) the inverse relationship between c and h can be observed in (3) for any fixed level of 'E':

$$E = \frac{hc}{\lambda}$$

Eq.3: Inverse Relationship Between c and h

Assume now a thought experiment to bring home the creative nature of light and why all of life would possibly emerge from it. Imagine light traveling at an infinite speed. If this were so then the inverse relationship between c and h would necessitate that h approach 0. In other words matter would be unable to form. Conversely if light were to slow down, to approach c, then h would progressively increase to some threshold where energy is able to sustain itself in quanta and matter would emerge. So finite speed of light creates quanta, which creates matter. Therefore all of matter emerges from light.

Note that such an interpretation of quanta suggests that quantum computation needs to be thought of differently than we currently think of it as. This was examined in 'Light-Based Interpretation of Quanta and its Implications on Quantum Computing' (Malik, 2020) and in The Emperor's Quantum Computer (Malik, 2018b). Further, every act of computation yields an output. It is proposed that genetic-type information is the output.

Further, in this view it can be seen that the big bang is nothing other than light slowing down and creating matter as a result of that (Malik, 2018a). So if all matter is an action of light then any universe arising is only a result of it. There are then likely some overarching properties that would be true of light and therefore also true of everything that was to arise in any such universe. It may even be that these properties would be fundamental and would determine structure of matter and everything that emerges out of it.

So, what can be inferred about the properties of light?

The speed of light is known to have implications on the experienced nature of reality. The finiteness, c, at 186,000 miles per second in a vacuum, creates an upper bound to the speed with which any physical object may travel. This also implies that objective reality will be experienced as a past, a present, and a future, from the point of view of that object (Einstein, 1995). These characteristics – a past, a present, and a future – can therefore be thought of as implicit in the nature of light and become part of objective reality because of the speed of light.

Further, it can be observed that c also creates a lower bound when inverted (1/c) being proportional to Planck's constant, h. 'h' as we know pegs the minimum amount of energy or quanta required for expression at the sub-atomic level (Isaacson, 2008). Planck's constant, h, therefore allows matter to form (Lorentz, 1925) and for the reality of nature with a past, present, and future, to also be progressively experienced as a phenomena of connection between seemingly independent islands of matter. This characteristic of 'connection' is therefore also proposed to be implicit in the nature of light and becomes part of objective reality because of the speed of light.

As suggested in The Fractal Organization (Malik, 2015) a 'present' equates to 'vitality' because in the present there is a working out of the play of forces where the most energetic, powerful, or 'vital' force will express itself over others. 'A 'past' equates to 'physicality' because all can be viewed as established reality, as defined by what the eye or other lenses of perception can see. Such lenses see what has already 'physically' been formed in time. A 'future' equates to 'mentality' since cause, or

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

seed, or direction, implies meaning that drives the emergence of phenomena.

These implicit characteristics – physical, vital, mental, connection – of the nature of light as experienced at the layer of reality set up by a finite speed of light may hence be summarized by (4). C_U refers to the speed of light of 186,000 miles per second, that has created the perceived nature of reality, U:

$$C_U: [Physical, Vital, Mental, Connection]$$

Eq. 4: Nature of Light at c

It is known however that at quantum levels the nature of reality is at least characterized by wave-particle duality. Light and matter may be experienced as both particles and waves (Feynman, 1985; De Broglie, 1929; Ekspong 2014). Such duality, as will be explored shortly is related to the notion of quantum. But for matter to be experienced as waves implies that 'h' has become a fraction of itself, $h_{fraction}$. This further implies that c must have become greater than itself, c_N, such that the inequality specified by (5) holds:

$$c_N > c_U$$

Eq. 5: Inequality of Speed of Light

Note that what is implied here is that there must be another nature of reality specified by N that is the result of a speed of light greater than 186,000 miles per second, just as there is a nature of reality specified by U that is the result of the speed of light being 186,000 miles per second. This is consistent with recent developments in physics with the notion of property spaces being separate from but influencing physical space as explored by Nobel Physicist Frank Wilczek (Wilczek, 2016).

It is also to be noted that in Perkowitz's recent treatment of today's breakthroughs in the science of light (Perkowitz, 2011) he suggests that the theory of relativity does not disallow particles already moving at speed c or greater.

It stands to reason that current instrumentation, experience, and normal modes of thinking having developed as a bi-product of the characteristics so created in the layer of reality U may be inadequate to access N without appropriate modification.

The notion of wave-particle duality already challenges the notion of normal thinking perhaps because particle-like phenomena may be viewed as a function of less than or equal to c motion, while wave-like phenomena may be viewed as a function of faster than c motion. That these may be happening simultaneously is reinforced by principles such as complementarity in which experimental observation may allow measurement of one or another but not of both (Whitaker, 2006).

III. Genetic-Type Information In Light

But then taking this trend of a possible increase in the speed of light to its limit, this will result in a speed of light of infinite miles per second. The question is, what is the nature of reality when light is traveling at infinite miles per second? And what kind of information would be created in that reality?

In any space-time continuum be it an area or volume, regardless of scale, light originating at any point will instantaneously have arrived at every other point. Hence light will have a full and immediate *presence* in that space-time continuum. Further, that light will *know* everything that is happening in that space-time instantaneously – that is know what is emerging, what is changing, what is diminishing, what may be connected to what, and so on - or have a quality of *knowledge*. It will connect every object in that space-time completely and therefore have a quality of connection or *harmony*. Finally nothing will be able to resist it or set up a separate reality that excludes it and hence it will have a quality of *power*.

These implicit characteristics of the nature of light as experienced at the layer of reality set up by light traveling infinitely fast may hence be summarized by (6), where c_∞ refers to the speed of light of ∞ miles per second, that has created the perceived nature of reality, ∞:

$$c_\infty: [Presence, Power, Knowledge, Harmony]$$

Eq. 6: Nature of Light at ∞

But it can also be noticed from (4) that 'physical' is related to presence, 'vital' is related to power, 'mental' is related to knowledge, and 'connection' is related to harmony.

The question then, is how do these apparent qualities at ∞ precipitate, translate into, or become the physical-vital-mental-connection based diversity experienced at U? This may be achieved through the intervention or action of a couple of mathematical transformations. First, the essential characteristics of Presence, Power, Knowledge, Harmony that it is posited exist at every point-instant by virtue of the ubiquity of light at ∞ will need to be expressed as sets with up to infinite elements. Such a precipitation is none other than an act of quantization since something implicit in the layer where light travels faster is collecting in 'quanta' to be expressed more materially in the layer where light travels slower. Second,

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

elements in these sets will need to combine together in potentially infinite ways to create a myriad of seeds or signatures that then become the source of the immense diversity experienced at U. This, similarly, is also an act of quantization or the action of a quantization-function. This suggests that all that is seen and experienced at U may be nothing other than 'information' or 'content' of light and as such that there are fundamental mathematical symmetries at play where everything at U is essentially the same thing that exists at ∞.

It may also be inferred that wherever wave-particle duality exists, it does so because of a more observable quantum-translation from one speed of light to another, through the device of quanta.

Assuming that the first transformation occurs at a layer of reality K where the speed of light is c_K, such that $c_U < c_K < c_\infty$, this may be expressed by (7):

$$c_K: [S_{Pr}, S_{Po}, S_K, S_H]$$

Eq. 7: Nature of Light at K

S_{Pr} signifies 'Set of Presence' and may have elements associated with the qualities of being present everywhere, or of creating a physical basis in or on which other functions or characteristics can manifest. Hence S_{Pr} may have elements as expressed by (8):

$$S_{Pr} \ni [Service, Diligence, Perseverance, Stability, \dots]$$

Eq. 8: Set of Presence

S_{Po} signifies 'Set of Power' and may have elements associated with the qualities of being powerful, or of the play of vitality and experimentation which creates all possibility. Hence S_{Po} may have elements as expressed by (9):

$$S_{Po} \ni [Power, Energy, Adventure, Experimentation, \dots]$$

Eq. 9: Set of Power

S_K signifies 'Set of Knowledge' and may have elements associated with the qualities of knowledge, or the search and codification of knowledge. Hence S_K may have elements as expressed by (10):

$$S_K \ni [Knowledge, Making\ of\ Laws, Spread\ of\ Knowl., \dots]$$

Eq. 10: Set of Knowledge

S_H signifies 'Set of Harmony' and may have elements associated with the qualities of harmony, or creating relationship and love. Hence S_H may have elements as expressed by (11):

$$S_H \ni [Harmony, Relationship, Love, Specialization, \dots]$$

Eq. 11: Set of Harmony

Equations (8 - 11) then, shed insight not only into the nature of reality in light's precipitation toward speed c, but also the type of information that may be generated in this precipitation. This information, it is proposed, is genetic-type information, and has a bearing on information that materializes in genes in the layer of reality where light travels at speed c.

Assuming that the second transformation occurs at a layer of reality N where the speed of light is c_N, such that $c_U < c_N < c_K < c_\infty$, this may be expressed by (12):

$$c_N: f(S_{Pr}\ x\ S_{Po}\ x\ S_K\ x\ S_H)$$

Eq. 12: Nature of Reality at N

The unique seeds are therefore a function, f, of some unique combination of the elements in the four sets S_{Pr}, S_{Po}, S_K, S_H. This also suggests the basis of vast genetic diversity, due to the functional variety of information available in layers of light antecedent to the layer traveling at c.

The relationship between the layers of light may be hypothesized by the following matrix (13):

$$\begin{bmatrix} c_\infty: [Pr, Po, K, H] \\ (\downarrow R_{C_K} = f(R_{C_\infty})) \\ c_K: [S_{Pr}, S_{Po}, S_K, S_H] \\ (\downarrow R_{C_N} = f(R_{C_K})) \\ c_N: f(S_{Pr}\ x\ S_{Po}\ x\ S_K\ x\ S_H) \\ (\downarrow R_{C_U} = f(R_{C_N})) \\ c_U: [P, V, M, C] \end{bmatrix}_{Light}$$

Eq. 13: Relationship Between Layers of Light

The matrix suggests a series of transformations leading from the ubiquitous nature of light implicit in a point – presence, power, knowledge, harmony - to the seeming diversity of matter observed at the layer of reality U which is fundamentally the same presence, power, knowledge, and harmony projected into another form of itself.

The first transformation is summarized by (14):

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

$$R_{C_K} = f(R_{C_\infty})$$

Eq. 14: First Transformation

This is suggesting that the reality at the layer specified by the speed of light c_K, R_{C_K} is a function of the reality at the layer specified by the speed of light c_∞. This transformation translates the essential nature of a point into the sets described by (8 – 11). Note that (14) is essentially a quantization-function, in that something of the reality of light existing at R_{C_∞}, is translated into reality experienced at R_{C_K}.

The second transformation is summarized by (15):

$$R_{C_N} = f(R_{C_K})$$

Eq. 15: Second Transformation

This is suggesting that the reality at the layer specified by the speed of light c_N, R_{C_N} is a function of the reality at the layer specified by the speed of light c_K. This transformation combines elements of the sets into unique seeds as suggested by (12). This transformation can also be thought of as the result of a quantization-function such that something of R_{C_K} is collected as unique seeds at R_{C_N}.

The third transformation is summarized by (16):

$$R_{C_U} = f(R_{C_N})$$

Eq. 16: Third Transformation

This is suggesting that the reality at the layer specified by the speed of light c_U, R_{C_U} is a function of the reality at the layer specified by the speed of light c_N. This transformation builds on the unique seeds suggested by (12) to create the diversity of U as specified by (4). This transformation is therefore also the result of a quantization-function such that the seed-aspect of R_{C_N} is translated into the immense diversity experienced at R_{C_U}.

In this framework the notion of wave-particle duality hence may become complementary block-field-wave-particle quadrality where block refers to phenomenon resident to ∞, field to phenomenon resident to N, wave to phenomenon resident to K, and particle to phenomenon resident to U. The essential translation from one level to the next is due to a series of quantization-functions, so that (13) essentially summarizes an algorithm for life (Malik at al., 2019), where an implicit quaternary basis of presence, power, knowledge, and harmony, sets up potentially infinite number of elements derived from sets of presence, power, knowledge, and harmony. The implication of this is that quantization, and in fact this genre of *quantum computation* that arbitrates a structure as summarized by (13) is fundamentally creative, resulting in the vast variety of genetic-type information. As such quantum computation should be thought of as a fundamentally creative process arbitrating abstract possibilities in Light into a rich variety of genetic-type information, that subsequently express itself in material existence.

Equation (13), with modification, can also be expressed as a light-based quantum computational model of genetics as in Eq (17). In this modification it is assumed that light is projected at zero-speed which would effectively create the opposite reality to light existing at infinite speed. Hence in (17) c_0 implies light at zero-speed, and D, W, I, and C imply Darkness, Weakness, Ignorance, and Chaos, the opposites of Presence, Power, Knowledge, and Harmony, respectively:

$$\begin{bmatrix} c_\infty: [Pr, Po, K, H] \\ (\downarrow R_{C_K} = f(R_{C_\infty})) \\ c_K: [S_{Pr}, S_{Po}, S_K, S_H] \\ (\downarrow R_{C_N} = f(R_{C_K})) \\ c_N: f(S_{Pr}\ x\ S_{Po}\ x\ S_K\ x\ S_H) \\ (\downarrow R_{C_U} = f(R_{C_N})) \\ c_U: [P, V, M, C] \\ \Uparrow \\ c_{0:\,[D,W,I,C]} \end{bmatrix}_{Light}$$

Eq. 17: Light-Based Quantum-Computational Model of Genetics

IV. Key Concepts of Genetics From the Light-Based Point Of View

1) The Origins of Genetics

The infinite information codified in Light is the origin of genetics. As discussed previously Light in its state where it travels infinitely fast, possesses characteristics of presence, power, knowledge, and harmony, and contains vast possibility within it. This vast possibility can be thought of as information, and its structure related to the four characteristics, as the basis of genetics.

2) The Light-Based Downward-Strand

All-possibility that exists in the reality of light traveling infinitely fast is progressively materialized through a mathematical arrangement by which the subtle-infinite becomes the astounding material-diversity experienced when light travels at c. The mathematical

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

process, by which this transformation takes place, creates the light-based downward-strand.

3) Subtle-Libraries of Pre-Genetic Information

The progressive materialization of light has been modeled by a series of mathematical transformations in the Sections II and III. These transformations take the vast amount of information existing where light travels infinitely fast, to effectively create a series of precipitated subtle-libraries also of practically infinite pre-genetic information. This series of subtle-libraries subsequently allows infinite material diversity to come into being.

4) Material-Fabric

Just as genetic information is housed in DNA in living cells, there has to exist some structure to house the proposed pre-genetic information that exists at a pre-cellular stage. It is proposed that such pre-genetic information is housed in a structure termed 'Material-Fabric'. This material-fabric exists at the interface or could be the interface between the antecedent quantum-layer and the reality that emerges at c_U.

5) Four-Base Logic-Encoding Ecosystems

Four-base logic-encoding ecosystems are imagined containing logic in the quantum-layer antecedent to the reality emerging at c_U. These four-base logic-encoding ecosystems can be subject to change in the interplay with the material layer and may be related to the processes of mutation.

6) Superposition in Genetics

In the process of quantum-computation, by which four-base logic-encoding ecosystems, the pre-genetic and the genetic libraries, can be altered, the different dynamics representative of realities created by light traveling at different speeds are always present. This presence exists in superposed fashion and the real-time quantum-computation determines which superposed possibility will manifest materially.

7) Entanglement in Genetics

The information inherent to a particular layer, created through light traveling at a different speed, generates libraries of possibility through a mathematical process. Due to different dynamics of space and time representative of the layer created by light at a particular speed, these libraries exist differently in an entangled state, therefore being subtly present or influencing layers of light traveling at a slower speed relative to that layer. Common DNA existing in every cell at the material layer may be thought of as a logical outcome of this process of antecedent-entanglement.

8) Heredity

The presiding or generally accessible four-base logic-encoding ecosystems may be thought of as the primary bases of heredity.

9) Constructive-Mutation

Constructive-mutation is imagined occurring when patterns of a largely obstinate nature at the material level are broken as a result of which other possibilities existing in the higher levels of the light-based downward-strand are allowed to manifest. Of necessity this means that an inherent process of integration is taking place since light is unifying with its deeper nature. Constructive-mutation is intimately tied to this notion of integration.

10) Destructive-Mutation

When obstinate or disintegrating patterns persist or are chosen destructive-mutation will result. In its essence this suggests that light is moving away from its essential unified reality more towards the reality typified by disaggregation, that can be imagined to exists were light to be projected at zero-speed as suggested by c_0 in Eq (17).

11) Interpretation of Matter

Matter is the result of a constant computation involving the existing material layer, the material-fabric, four-base logic-encoding ecosystems, and the antecedent light layers. This means that matter can and will change as the interplay between these layers changes.

12) Evolution and the Possibilities of Genetics

Evolution is a process by which pre-genetic possibilities in antecedent layers of light materialize due to the existence of the right material conditions. This will also imply that the pre-genetic material existing in the four-base logic-encoding ecosystems will change.

13) Post-Genetic Code

In its possibilities of materialization information in light may house itself in subtle-libraries generated at various layers of light, four-base logic-encoding

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

ecosystems at the quantum-level, the material-fabric, or in genetic code in living cells. But it is also possible through the process of evolution that the "structure" housing further possibilities in light may take on a hybrid form that may be referred to as Post-Genetic Code.

V. Summary and Further Research

The light-based quantum-computational model of genetics described in this paper suggests several avenues of further research.

In this model there are multiple layers of light that house different kinds of fourfold information. The model is subject to a persistent quantum computation which creates genetic-type information as its output. In this model the origin of genetics is seen as being the antecedent layers of light traveling faster than c. Such layers of light can mathematically be structured as non-physical property-spaces that impact physical reality. The question is whether such a configuration of light-based property-spaces are valid in terms of thinking about genetics, and further what other kinds of property-space models may be valid as a basis of comparison?

The precipitation of light is suggested to have a bearing on the downward strand of DNA. Subtle-libraries of genetic-type information are suggested to be inputs in a process of quantum computation that also have a bearing on how genetic information materializes. In terms of further research, is there indeed such a precipitation in the existing conception of the downward-strand of regular DNA? Further, is there some relationship between the four nitrogenous bases in DNA of guanine, adenine, cytosine and thymine and the four properties envisioned to be implicit in light?

For such a computational model to be viable there needs to be another housing structure for genetic-type information. It is proposed that there is some kind of 'material-fabric' at the quantum level where light precipitates to light at speed c. Is this indeed the case?

Further, as a result of being at the quantum-level, phenomenon such as superposition and entanglement may also influence genetic information. Is the model of superposition and entanglement proposed briefly in this paper a possible way in which these quantum phenomenon affect genetics?

Persistence of information in such a material-fabric is the basis of heredity. Constructive mutation is due to interaction with layers of light traveling faster than c where the root of all function is proposed to exist. Destructive mutation may be perceived as being due to interaction with light existing at zero-speed, where disaggregation increases. These views of heredity and mutation need to be elaborated further.

If light is indeed the origin of genetics and if light-based processes such have been briefly discussed in this paper are valid, then what are implications for non-invasively and positively influencing processes of genetics? Experiments to begin to check the validity of this possibility can be created.

References

1. De Broglie, L. 1929. The Wave Nature of the Electron. Nobel Lecture.
2. Einstein, A. 1995. Relativity: The Special and General Theory. New York: Broadway Books.
3. Ekspong. 2014. "The Dual Nature of Light as Reflected in the Nobel Archives". *Nobelprize.org*. Nobel Media AB 2014. Web. 15 Oct 2016.
4. Feynman, RP. 1985. QED The Strange Theory of Light and Matter. New Jersey: Princeton University Press
5. Holland, P. 1995. The Quantum Theory of Motion: An Account of the de Broglie-Bohm Causal Interpretation of Quantum Mechanics. Cambridge: Cambridge University Press.
6. Isaacson, W. 2008. Einstein: His Life and Universe. Simon and Schuster. New York.
7. Lloyd, S. 2007. Programming the Universe: A Quantum Computer Scientist Takes On the Cosmos. New York: Vintage
8. Lorentz, H.A. 1925. The Science of Nature. Vol. 25, p 1008. Springer
9. Malik, P. 2015. The Fractal Organization. New Delhi: Sage Publications
10. Malik,P., Pretorius, L., Winzker, D. 2017. Qualified Determinism in Emergent-Technology Complex Adaptive Systems. Conference Proceedings IEEE TEMSON 2017.
11. Malik, P. 2018a. Cosmology of Light. Google Books

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

12. Malik, P. 2018b. The Emperor's Quantum Computer. Google Books.
13. Malik, P. 2019. The Origin and Possibilities of Genetics. Google Books.
14. Malik, P. Pretorius, L. 2019. An Algorithm for the Emergence of Life Based on a Multi-Layered Symmetry-Based Model of Light. 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC). 10.1109/CCWC.2019.8666554
15. Malik, P. "Light-Based Interpretation of Quanta and its Implications on Quantum Computing," 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2020, pp. 0719-0726, doi: 10.1109/CCWC47524.2020.9031279.
16. Perkowitz, S. 2011. Slow Light. London: Imperial College Press
17. Ridley, M. Genome: The Autobiography of a Species in 23 Chapters.. Great Britain: Fourth Estate.
18. Whitaker, A. 2006. Einstein, Bohr and the Quantum Dilemma: From Quantum Theory to Quantum Information. Cambridge: Cambridge University Press
19. Wilczek, F. 2016. A Beautiful Question: Finding Nature's Deep Design. New York: Penguin Books

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Barriers and Drivers to Adoption of Cloud Infrastructure Services: A Security Perspective

Michael Townsend
Departmetn of Information Technology
Georgia Southern University
Statesboro, USA
mtownsend@georgiasouthern.edu

Hayden Wimmer*
Department of Information Technology
Georgia Southern University
Statesboro, USA
hwimmer@georgiasouthern.edu

Jie Du
School of Computing and Information Systems
Grand Valley State University
Allendale, USA
dujie@gvsu.edu

***Abstract*—This study investigates the barriers and drivers to cloud infrastructure adoption as perceived by Information Technology Professionals. If the barriers to adoption of cloud computing technologies exist in small then the identification of these barriers will help organizations implement cloud technologies into their operating environments. This study will focus on the adoption of Infrastructure-as-a-Service technologies and not Software-as-a-Service technologies. Much of the literature on this topic revolves around security and the perception that it is less secure. While there are security issues to be addressed, many of the barriers can be overcome through an understanding of the cloud technology. A survey was conducted to measure the perception and readiness of Information Technology professionals and business decision makers. The results of the survey show that the major barriers to adoption are not always security related and that some benefits may also be found. Perceived security is found to be a significant driver for adoption of IaaS.**

Keywords—cloud computing, security, adoption, barriers

I. INTRODUCTION

Managing information technologies can be a complex issue and one that can be done in many different ways. There are traditional on premise datacenters that provide local computing resources and then there are large, regional datacenters available that can accommodate most any workload an organization would need to acquire. With all of the benefits that these large datacenters can provide, and the most significant players in the market being household names like Amazon, Microsoft and Google; why are not more organizations taking advantage of the easily accessible and robust computing environment?

This study will investigate the barriers and the drivers to cloud infrastructure adoption as perceived by IT professionals in the United States. The focus of this study will be on the adoption of Infrastructure-as-a-Service (IaaS) technologies and not Platform-as-a-Service technologies like databases or Software-as-a-Service technologies like Google Gmail and Microsoft Office365. A survey will be conducted to measure the perception and readiness of Information Technology professionals to adopt cloud technologies over an on premise datacenter infrastructure option.

This work will contribute to the body of Information Technology knowledge by providing a specific point of view for adopting cloud computing infrastructure technologies. Much of the literature written on this topic is broader and not focused on specific infrastructure technologies and the determinants to adopting cloud infrastructure technologies.

This paper starts with a discussion of research into cloud computing and then focuses on the cloud computing security implications, perceived security, and trust. A description of the research model and methodology we applied to help answer our research question follows. Finally, it concludes with a presentation of the data analysis and a discussion of their implications.

II. LITERATURE REVIEW

The National Institute of Standards and Technology (NIST) defines cloud computing as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction" [1]. Cloud computing is becoming a more pervasive and essential part of IT departments in business, government, and non-profit sectors. The impact of virtualization and distributed computing has created an opportunity for cloud vendors to capitalize on excess computing infrastructure [2]. Infrastructure-as-a-Service is the model where the consumer provisions the processing, storage, and networking to include operating system and any application software. The consumer does not have access to the base hardware architecture that is physically located in the provider's datacenter.

Armbrust et al. [3] provided an overview and advocated for the use of cloud computing through discussions related to utility computing, Internet datacenters, and distributed system economics. Armbrust et al. included information and advocacy for the use of cloud computing and gave example use cases for Amazon EC2, Microsoft Azure, and Google AppEngine. They include a Top 10 obstacles and opportunities for cloud computing related to uptime, security, networking bottlenecks, scaling, licensing and unpredictability in performance and software integrity. The third topic discussed in [3] is on the topic of Data Confidentiality and Auditability. Since cloud offerings are essentially public networks, they are more prone to attack than a private, on premise cloud. "We believe that there are no fundamental obstacles to making a cloud-computing environment as secure as the vast majority of in-house IT environments, and that many of the obstacles can be overcome immediately with well understood technologies such as encrypted storage, Virtual Local Area Networks, and network

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

middle boxes (e.g. firewalls, packet filters)" [3]. They proposed that encrypting data before placing it in a Cloud may be even more secure than unencrypted data in a local datacenter [3].

A. Cloud Security Implications

Almorsy, Grundy, and Muller [4] discussed security implications for Infrastructure-as-a-Service to include virtual machine security such as the operating system which is the responsibility of the cloud consumer. Securing the virtual machine images repository as a compromised image can result in the spread of malicious software to numerous hosts. Security of images is the burden of the cloud provider. Virtual network security is important because multiple organizations could share the same server and virtual network and Domain Name Service and other IP protocols could be exploited. The virtual boundaries of each virtual machine on a given host server should be isolated and it is the responsibility of the cloud provider. These virtual machines share the same processor, memory, network card, and input/output stream. Lastly, the hypervisor is what virtualizes and maps virtual resources to physical resources on the host server. "Any compromise of the hypervisor violates the security of the VMs because all VMs operations become trace unencrypted" [4].

More technical sources cite the difficulty of digital investigations on cloud infrastructure. Almulla, Iraqi, and Jones [2] discussed the infrastructure and services required to serve cloud applications to end users. Investigations with the cloud and on the cloud are discussed with importance of guidelines in relation to assist the investigator in a cloud scenario [2]. Subashini [5] made some valid points regarding security issues in IaaS and the responsibilities of both the provider and the consumer as to how they differ between cloud service models. Amazon Elastic Compute Cloud (EC2) infrastructure as a service model includes vendor responsibility for security controls for the physical, environmental, and virtualization layers. The customer is responsible for the security of the operating system, applications, and data [5].

El-Gazzar, Hustad, and Olsen [6] found that two of the most important factors that drive cloud computing adoption are security and governance policies designed to monitor the performance of the cloud service provider. Using structural equation modeling and artificial neural network modeling, Priyadarshinee, Raut, Jha, and Gardas [7] found that trust, perceived IT security risk, management style, and technology innovation were identified to have an important effect on cloud adoption. Technical barriers to adoption were researched by [8] and they found that large scale outages that impacted availability, lack of portability, and integration with current systems impacted adoption. Furthermore, migration complexity, privacy and legal concerns, and security all play a role in whether a business will adopt [8].

In a localized study of Tunisian firms Hadhri [9] posited that adoption depended on that size of the firm, the savings that cloud computing could provide, complexity, and competitive pressure from other businesses. An interesting find is that one component that was not important was the relative IT skills of the employees.

"Cloud computing involves potentially greater exposure to security threats and privacy breaches, especially when the cloud is based on the Internet rather than an organization's own internal network" [10]. Despite the security concerns cloud computing is a desirable solution for many companies. New small businesses can get started without having to pay for hardware and operating systems up front. Current small business can migrate existing workloads to the cloud with very little downtime and enjoy the benefits of paying for what you use.

Xiao and Xiao [11] discussed confidentiality, integrity, availability, accountability, and privacy that are the most representative security and privacy attributes in cloud computing. Along with vulnerabilities like co-residency (multi-tenancy) and loss of physical control in each attribute, defense strategies and suggestions were also discussed such as co-residency detection and trusted computing platform. Privacy and security issues in cloud computing was studied by Kshetri [12] and found that "due primarily to concerns related to security, privacy and confidentiality critics have argued that its perceived costs may outweigh the benefits" [12]. These findings help solidify the hypotheses that security is indeed a primary barrier to adoption of cloud computing.

Kahn and Al-Yasiri [13] investigated challenges associated with cloud adoption among small and medium businesses in the United Kingdom. They identified, through interviews, four main challenges encountered by small and medium business, lack of internal staff expertise, concern over proprietary data and its safety, prohibition of classified data being placed on a shared or outsourced infrastructure, the last and most predominate problem was about Service Level Agreements (SLA) - regulatory and service terms in cloud contracts and the lack of familiarity with the cloud concepts discussed in the SLA.

B. Perceived Security and Trust

Sun, Change, Sun, and Wang [14] presented that high security is one the major obstacles to adoption of cloud computing as a utility. The vulnerabilities to accessibility, virtualization, and web applications pose many security challenges. Security issues in cloud computing that can be divided into six subcategories, (a) how to provide safety mechanisms, (b) how to keep data confidential, (c) avoiding malicious insiders, (d) avoiding service hijacking, (e) managing multi-tenancy in virtualized environment where more than one entity resides on a host server owned by the service provider, and (f) developing "appropriate law and legal jurisdiction, so that users have a chain against their providers if need" [14].

According to [14] privacy is "the ability of an individual or group to seclude themselves or information about themselves and thereby reveal themselves selectively." Privacy issues can be divided into four subcategories, (a) how to allow users to keep control of their data when it is stored and processed in the cloud, (b) how to maintain the integrity of the user's data when it is replicated to other jurisdictional areas and avoid data loss and unauthorized modification, (c) defining which party is responsible for ensuring legal requirements for personal information are met, and (d) the identification and vetting of sub-contractors that might have access to the system that contains user data.

Lastly, Sun, Change, Sun, and Wang [14] addressed the issue of trust. "Trust is viewed as a measurable belief that utilizes experience to make trustworthy decisions" [14]. Trust can be built by restricting or blocking malicious entities from accessing or infiltrating the service provider's systems. The four subcategories of trust are, (a) how to define and evaluate trust in a specialized cloud computing environment, (b) incident response and handling, (c) providing a different level of security based on trust level chosen by the end user, and (d) how to respond in time to changes in trust relationships when user requirements change.

"Trust and trustworthiness are foundations on security" [15]. The basis of trust relationships and trust formation can dramatically affect the underlying security of any system [15]. These trust assumptions can affect the analysis of security requirements [16, 17]. Elahi and Yu examined the trust trade-offs in security requirements [18]. Aslam, Bouget, and Raza proposed a solution to preserve security and trust in virtual machine migrations [19].

"Diffusion is a critical process concerning the success of innovations" [20]. This term relates directly to the study discussed in this paper in that the perceived security and trust of cloud technologies and that subsequent adoption is a long and slow process. The adoption rate of cloud technologies is at odds in some cases. Stieninger and Nedbal [20] found that cloud computing supporters grew from 28 to 35 percent, and the group of opponents also grew from 38 to 44 percent from 2011 (sample size n=411) to 2012 (sample size n=436). The fact that these two groups have opposing views and are also growing in size is much of the motivation for this research. What is the barrier to adoption for the holders of the opposing viewpoint?

There are two main organizations working toward a more secure cloud computing experience and solutions and standards. The Cloud Security Alliance (CSA) and the Open Web Application Security Project (OWASP). CSA fosters discussion among solution providers, non-profit entities, and individuals for recommended practice for securing cloud computing. CSA also observes existing security framework standards and goes beyond that by creating a new framework. "CSA bridges the gap between common industry and government standards by offering its own framework of consolidated standards" [21]. OWASP curates a list of top cloud vulnerabilities to cloud based Software-as-a-Service models which is updated based on the threat landscape present at the time [22].

III. METHODOLOGY

A. Research Model

Our research model (see Fig. 1) has six constructs: perceived security [23], trust [24], risk [25], system quality [26], perceived benefits [26], and attitude [27]. The dependent variable in the research model is the subjects' self-reported attitude toward IaaS (ATT). Five independent variables are hypothesized: the perceived security of IaaS, the trust to adopting IaaS, the perceived risks of adopting IaaS, the system quality of IaaS, and the perceived benefits of adopting IaaS.

- H1 – Perceived Security (SEC) of IaaS is positively related to the attitude toward IaaS.
- H2 – Trust (TRU) is positively related to the attitude toward IaaS.
- H3 – Perceived risk (RISK) of adopting IaaS is negatively related to the attitude toward IaaS.
- H4 – System quality (QUA) of IaaS is positively related to the attitude toward IaaS.
- H5 – Perceived benefits (BEN) of adopting IaaS are positively related to the attitude toward IaaS.

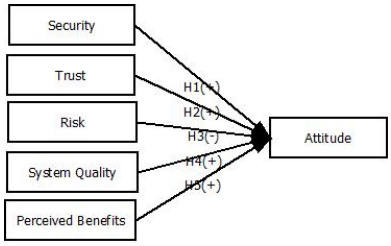


Fig. 1 Research Model

B. Survery Administration

The items in the survey focused on six constructs including five IVs and one dependent variable. All items (see the appendix) are anchored on 7-point Likert scales, with the following weights on each response, Strongly Disagree = 1, Somewhat Disagree = 2, Disagree = 3, Neutral = 4, Agree = 5, Somewhat Agree = 6, and Strongly Agree = 7. Survey participants were colleagues and peers of one of the authors that were encouraged to share the survey. An email was sent to the participants to explain the study and inform them the time estimated to complete along with other information required by the university. A survey service from Qualtrics was used to administer the survey questions. The survey consisted of 8 demographic questions and 40 simple technical questions with a target completion time of less than 30 minutes as it is difficult to get anyone to commit much time to a survey.

The survey recipient list consisted of 40 curated emails from personal contacts and professionals in the area. 27 responses were received and used in this study. Table 1 summaries the demographics of the sample.

TABLE 1. SURVEY DEMOGRAPHIC PROFILE

Demographic Category	Percentage
Males	92%
Females	8%
Average Experience	17 years
Age Representation	25-34: 37% 35-44: 30% 45-54: 30% 55-64: 3%
Education	Some College: 8% 2-year degree: 8% 4-year degree 38% Professional Degree: 42% Doctorate: 4%

IV. DATA ANALYSIS

A two-phased analysis was conducted to investigate the barriers and the drivers to adoption of IaaS. In Phase 1, we conducted an empirical study on the survey questions. We compared each pair of questions on IaaS and on-premise using paired *t*-tests. In Phase 2, we conducted a two-step analysis to examine the effects of the key constructs on the attitude dependent variable. First, an exploratory factor analysis (EFA) was done to extract the factors (latent variables) to validate our model constructs. Second, a multiple regression analysis was conducted using the SPSS calculated factor scores. The dependent variable was regressed on the five IVs to determine the main effects.

A. Phase 1: Comparing IaaS and On Premise

We compared the 20 pairs of questions on IaaS and on premise (see Fig. 2 - 7). Of the 20 question sets presented to the respondents, 5 were measured to have a significant difference (see Table 5). IaaS is perceived to have higher availability than on premise ($p = 0.012$) (see Fig. 5). Perception exists that IaaS provides a significant boost in uptime. A strong perception exists that the respondents believe that IaaS services ability to scale are greater than on premise services ($p=0.000$) (see Fig. 6). IaaS is also perceived to have greater performance than on premise ($p = 0.024$) (see Fig. 4). The respondents felt that there was more reason to believe that poor performance would be found on premise and not in IaaS.

TABLE 2. FIVE SETS OF QUESTIONS AND THE P-VALUE

No.	Question	p-value
1	I trust Infrastructure as a Service (IaaS) because it has our best interests in mind.	0.003
	I trust On-Premise Service because it has our best interests in mind.	
2	It is probable that Infrastructure as a Service (IaaS) would frustrate me because of its poor performance.	0.024
	It is probable that On Premise Service would frustrate me because of its poor performance.	
3	Infrastructure as a Service (IaaS) is user friendly.	0.066
	On Premise Service is user friendly.	
4	Infrastructure as a Service (IaaS) provides high availability.	0.012
	On Premise Service provides high availability.	
5	Infrastructure as a Service (IaaS) helps the organization to respond more quickly to change.	0.000
	On Premise Service helps the organization to respond more quickly to change.	

On premise is found to have a better customer service ($p = 0.003$) (see Fig. 3). There is concern that cloud service providers and their ability to keep customer best interests in mind. There is a perception for the respondents that on premise is more user friendly ($p=0.066$, see Fig. 5). Cloud providers have expansive and numerous catalog services and applications designed to meet the needs of any customer and these offerings are always changing. On premise services are not as dynamic and are more static making it easier to use and not have to contend with an ever changing landscape of services.

Fig. 2 Security Comparison

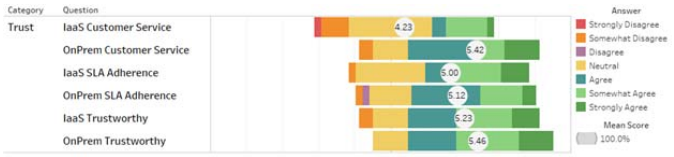


Fig. 3 Trust Comparison

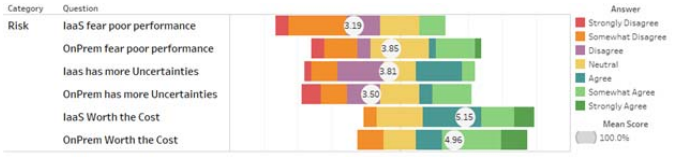


Fig. 4 Risk Comparison

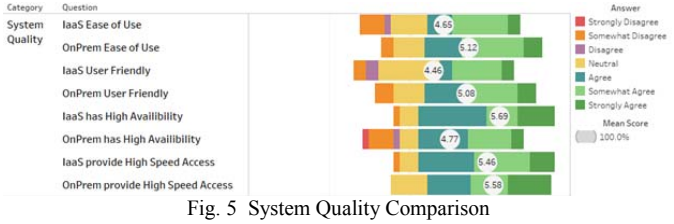


Fig. 5 System Quality Comparison

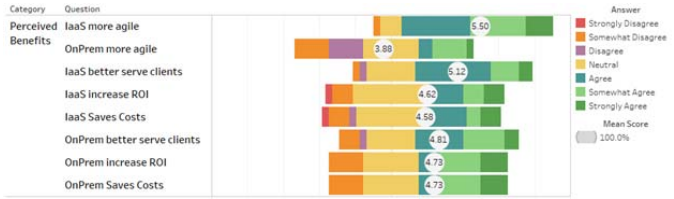


Fig. 6 Perceived Benefits Comparison

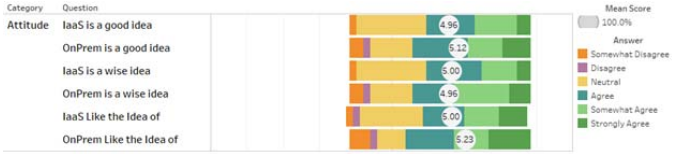


Fig. 7 Attitude Comparison

B. Phase 2:Determinants to Adoption of IaaS

1) Construct Validity and Reliability

We conducted the factor analysis (using primary axis analysis) on the data set to extract the factors that influence IT professionals' attitude toward IaaS. We use 0.5 as the recommended threshold [28]. Two rounds were run before we arrived at a set of factors loading at 0.5 or above. In the first run, four survey questions (TRU1, TRU2, RISK1, and BEN3) having a factor loading lower than 0.5 were removed from further consideration.

The results of the second round of EFA resulted in six factors being extracted from the data: ATT, SEC, BEN, QUA12, QUA34, and RISK. Note that the EFA resulted in the removal of the trust (TRU) factor from the original proposed model due to low factor loadings. TRU3 and BEN4 were loaded together with ATT1, ATT2, and ATT3. The EFA also resulted in the splitting of the original system quality factor (QUA) into two factors: QUA12, and QUA34. These unexpected results will be addressed in the discussion section later in this paper.

Cronbach Alpha coefficient was used to test the reliability of the items. The acceptable value of Cronbach Alpha should be at least 0.70 [29]. Table 3 summarizes the factor loadings and Cronbach Alpha values for each item. The factor loadings for all items are greater 0.5 and the Cronbach Alpha values for all factors are greater than 0.7, allowing us to proceed with our regression analysis and hypothesis testing.

TABLE 3. FACTOR LOADINGS AND CRONBACH ALPHA

Construct	Item	Factor loadings	Cronbach Alpha
ATT			0.928
	ATT1	.965	
	ATT2	.899	
	ATT3	.971	
	TRU3	.665	
	BEN4	.688	
SEC			0.953
	SEC1	.938	
	SEC2	.922	
	SEC3	.923	
BEN			0.987
	BEN1	.915	
	BEN2	.931	
QUA12			0.901
	QUA1	.906	
	QUA2	.650	
QUA34			0.888
	QUA3	.897	
	QUA4	.596	
RISK			0.766
	RISK1	-.708	
	RISK2	-.946	

2) Hypotheses Testing

To test the hypotheses, a multiple regression analysis was conducted using SPSS (see Table 4). The latent variable, perceived security has a significant coefficient as expected (p=0.095). The perceive security has a significant, positive effect on IT professionals' attitude toward IaaS. Thus, H1 was supported. Other variables were not significant. Therefore H2 - H5 were not supported.

TABLE 4. REGRESSION RESULTS

Variables	Coefficient
SEC	.333*
BEN	.215
QUA12	.059
RISK	.147
QUA34	.120
R^2	.408
Adjusted R^2	.267

V. DISCUSSION

This study identified positive perceptions about IaaS technologies that are also good reasons to adopt the technology. Subjects' responses indicate that IaaS is perceived to have higher availability and agility. IaaS is also perceived to have greater performance. On premise is found to have a more user friendly attribute and better customer service. Customer service and user friendliness are barriers to IaaS adoption. There is concern that cloud service providers and their ability to keep customer best interests in mind.

Subjects' responses also indicate that perceived security is a significant determinant of the attitude toward adoption of IaaS. Support for H1 indicates that IT professionals who believe that IaaS has adequate security have a more positive attitude toward adoption of IaaS.

A. Construct Loading

Besides the three items ATT1, ATT2, and ATT3, two other items (TRU3 and BEN4) were also loaded in the construct of ATT. TRU3 (I trust IaaS because it has our best interests in mind.) is phrased differently than other items in the TRU construct. TRU3 emphasizes more on the attitude or behavior other than the customer service. That might explain why TRU3 was loaded as ATT. Different from other perceived benefits questions focusing on specific benefits, BEN4 (IaaS helps the organization provide better services to its clients) leans toward IT professional's attitude toward IaaS rather than the perceived benefits. The question as to how people perceive the benefits and trust of using IaaS should be explored in the future.

B. Splitting a Model Construct

Our analysis also resulted in the splitting of the system quality factor (QUA) into two separate factors (QUA12, QUA34) based on the EFA results. A simple look at the questions gives insight into why this may have been necessary (Table 5). The first two items relate to the perceived quality of IaaS while the second two items focus on the performance of IaaS. Given these differences, it seems at least logical that the system quality factor needs to be split. The question as to how people perceive the system quality of IaaS is one that should be explored in the future.

TABLE 5. SYSTEM QUALITY SURVEY ITEMS

Item	Question
QUA1	Infrastructure as a Service (IaaS) is easy to use.
QUA2	Infrastructure as a Service (IaaS) is user friendly
QUA3	Infrastructure as a Service (IaaS) provides high availability
QUA4	Infrastructure as a Service (IaaS) provides high-speed information access

C. Limitation

There are two limitations that must be acknowledged regarding this research. One limitation of the research is that this was a small sample size and a larger sample size would be more desirable. Better results may have been obtained if the sample size had been greater. The second is that, while the survey was built from prior surveys in cloud computing security, some of the issues with factor loadings could be caused by weaknesses in the survey itself.

D. Implications for Research and Practice

There are at least two implications of these findings for the research community. First, perceived security is a significant determinant to adoption of IaaS. Instead of being a barrier, perceived security is a driver to adoption of IaaS. The limited number of significant factors in the model could be an indicator that better models are needed.

Implications for practice focus around informing professionals about the barriers to adoption of IaaS so that they can make better decisions when deploying applications. Customer service and user friendliness are barriers to IaaS adoption. This research looks to inform practitioners in Information Technology about options they can use to better deploy applications that allow workers to focus on customer support and not facility and server maintenance. One of the best indicators for adoption is the decision maker's knowledge of the technology.

VI. CONCLUSION

With more companies offering cloud computing services through the use of remote datacenters located all over the world, it is a compelling question as to what factors exist that impact the adoption of infrastructure cloud technologies. This study identified positive perceptions about IaaS technologies that are also good reasons to adopt the technology. Customer service and user friendliness were found to be barriers to IaaS adoption. Perceived security was identified to be a significant factor to positively impact IT professional's attitude toward adoption of IaaS.

Future work could focus on cost as compared to security. Many of the works do not mention cost as a major barrier but it must exist in some organizations. Future work could also center on the concept of diffusion and acceptance of cloud computing in general. Additionally, more work can be done on specific industries like higher education and K-12 education as well as individual business sectors.

REFERENCES

[1] P. Mell and T. Grance. (2011). *The NIST definition of cloud computing*.

[2] S. Almulla, Y. Iraqi, and A. Jones, "Cloud forensics: A research perspective," in *2013 9th International Conference on Innovations in Information Technology (IIT)*, 2013, pp. 66-71.

[3] Armbrust, Michael, A. Fox, Armando, Griffith, Rean, Joseph, A. D, R. Katz, R. H, A. Konwinski, Andrew, G. Lee, Gunho, Patterson, D. A, Rabkin, Ariel, Stoica, and Matei, "Above the Clouds: A Berkeley View of Cloud Computing," 2009.

[4] M. Almorsy, J. Grundy, and I. Mueller, *An analysis of the cloud computing security problem*, 2010.

[5] S. Subashini and V. Kavitha, "A Survey on Security Issues in Service Delivery Models of Cloud Computing," *The Journal of Network and Computer Applications, Elsevier,* vol. 35, pp. 1-11, 2011.

[6] R. El-Gazzar, E. Hustad, and D. H. Olsen, "Understanding cloud computing adoption issues: A Delphi study approach," *Journal of Systems and Software,* vol. 118, pp. 64-84, 2016/08/01/ 2016.

[7] P. Priyadarshinee, R. Raut, M. Jha, and B. Gardas, "Understanding and predicting the determinants of cloud computing adoption: A two staged hybrid SEM - Neural Networks Approach," *Computers in Human Behavior,* vol. 76, 2017.

[8] N. Phaphoom, X. Wang, S. Samuel, S. Helmer, and P. Abrahamsson, "A Survey Study on Major Technical Barriers Affecting the Decision to Adopt Cloud Services," *Journal of Systems and Software,* vol. 103, 2015.

[9] W. Hadhri, T. Maherzi, and A. B. Youssef, "E-Skills and the Adoption of Cloud Computing," *Thunderbird International Business Review,* vol. 59, pp. 635-645, 2017.

[10] M. Taylor, J. Haggerty, D. Gresty, and D. Lamb, "Forensic investigation of cloud computing systems," *Network Security,* vol. 2011, pp. 4-10, 2011/03/01/ 2011.

[11] Z. Xiao and Y. Xiao, "Security and Privacy in Cloud Computing," *Communications Surveys & Tutorials, IEEE,* vol. 15, pp. 843-859, 2013.

[12] N. Kshetri, "Privacy and security issues in cloud computing: The role of institutions and institutional evolution," *Telecommunications Policy,* vol. 37, pp. 372–386, 2013.

[13] N. Khan and A. Al-Yasiri, "Framework for Cloud Computing Adoption: A Roadmap for Smes to Cloud Migration," *International Journal on Cloud Computing: Services and Architecture,* vol. 5, pp. 01-15, 2015.

[14] D. Sun, G. Chang, L. Sun, and X. Wang, "Surveying and Analyzing Security, Privacy and Trust Issues in Cloud Computing Environments," *Procedia Engineering,* vol. 15, pp. 2852-2856, 2011/01/01/ 2011.

[15] J. Viega, T. Kohno, and B. Potter, "Trust (and mistrust) in secure applications," *Commun. ACM,* vol. 44, pp. 31–36, 2001.

[16] C. B. Haley, R. C. Laney, J. D. Moffett, and B. Nuseibeh, "The effect of trust assumptions on the elaboration of security requirements," in *Proceedings. 12th IEEE International Requirements Engineering Conference, 2004.*, 2004, pp. 102-111.

[17] S. Forti, G.-L. Ferrari, and A. Brogi, "Secure Cloud-Edge Deployments, with Trust," *Future Generation Computer Systems,* vol. 102, pp. 775-788, 2020/01/01/ 2020.

[18] G. Elahi and E. Yu, "Trust Trade-off Analysis for Security Requirements Engineering," in *2009 17th IEEE International Requirements Engineering Conference*, 2009, pp. 243-248.

[19] M. Aslam, S. Bouget, and S. Raza, "Security and trust preserving inter- and intra-cloud VM migrations," *International Journal of Network Management,* pp. 1-19, 2020.

[20] M. Stieninger and D. Nedbal, *Diffusion and Acceptance of Cloud Computing in SMEs: Towards a Valence Model of Relevant Factors*, 2014.

[21] C. D. Giulio, R. Sprabery, C. Kamhoua, K. Kwiat, R. H. Campbell, and M. N. Bashir, "Cloud Standards in Comparison: Are New Security Frameworks Improving Cloud Security?," in *2017 IEEE 10th International Conference on Cloud Computing (CLOUD)*, 2017, pp. 50-57.

[22] OWASP. (2013, August 19). *The OWASP Top Ten*. Available: https://owasp.org/www-pdf-archive/OWASP_Top_10_-_2013.pdf

[23] C. Flavian and M. Guinalíu, "Consumer trust, perceived security and privacy policy: Three basic elements of loyalty to a web site," *Industrial Management and Data Systems,* vol. 106, pp. 601-620, 2006.

[24] P. Pavlou, "Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model," *International Journal of Electronic Commerce,* vol. 7, pp. 101-134, 2003.

[25] I. Im, Y. Kim, and H. Han, "The effects of perceived risk and technology type on users' acceptance of technologies," *Information & Management,* vol. 45, pp. 1-9, 2008.

[26] Y.-S. Wang, H.-Y. Wang, and D. Shee, "Measuring e-learning systems success in an organizational context: Scale development and validation," *Computers in Human Behavior,* vol. 23, pp. 1792-1808, 2007.

[27] V. Venkatesh, M. Morris, G. Davis, and F. Davis, "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly,* vol. 27, pp. 425-478, 2003.

[28] J. F. Hair, R. E. Anderson, R. L. Tatham, and W. C. Black, *Multivariate Data Analysis (5th ed.)*. Upper Saddle River, NJ: Prentice Hall, Inc., 1998.

[29] J. Nunnally and I. H. Bernstein, *Psychometric Theory*. New York: McGraw-Hill Higher Ed. , 1994.

Appendix

Demographics
1. What is your age?
2. What is your race?
3. What is your gender?
4. What is your highest level of education?
5. What is your title in your current position?
6. How many years of experience do you have in your field?
7. Do you have experience with Infrastructure as a Service (IaaS)?
8. Do you have experience with On-Premise infrastructure?

Security - Infrastructure as a Service (IaaS) [23]
IaaS1 (SEC1). I think that Infrastructure as a Service (IaaS) has mechanisms to ensure safe transmission of user information.
IaaS2 (SEC2). I think that Infrastructure as a Service (IaaS) has sufficient technical capacity to ensure data will not be intercepted by unauthorized parties.
IaaS3 (SEC3). I think that Infrastructure as a Service (IaaS) has sufficient technical capacity to ensure data will not be modified by unauthorized parties.

Trust - Infrastructure as a Service (IaaS) [24]
IaaS4 (TRU1). Infrastructure as a Service (IaaS) is trustworthy.
IaaS5 (TRU2). Infrastructure as a Service (IaaS) keeps promises and commitments.
IaaS6 (TRU3). I trust Infrastructure as a Service (IaaS) because it has our best interests in mind.

Risk - Infrastructure as a Service (IaaS) [25]
IaaS7 (RISK1). It is probable that Infrastructure as a Service (IaaS) would be worth its cost.
IaaS8 (RISK2). It is probable that Infrastructure as a Service (IaaS) would frustrate me because of its poor performance.
IaaS10 (RISK3). Comparing with other technologies, using Infrastructure as a Service (IaaS) has more uncertainties.

System Quality - Infrastructure as a Service (IaaS) [26]
IaaS11 (QUA1). Infrastructure as a Service (IaaS) is easy to use.
IaaS12 (QUA2). Infrastructure as a Service (IaaS) is user friendly.
IaaS13 (QUA3). Infrastructure as a Service (IaaS) provides high availability.
IaaS14 (QUA4). Infrastructure as a Service (IaaS) provides high-speed information access.

Perceived Benefits - Infrastructure as a Service (IaaS) [26]
IaaS15 (BEN1). Infrastructure as a Service (IaaS) helps the organization save costs.
IaaS16 (BEN2). Infrastructure as a Service (IaaS) helps the organization to increase return on investment.
IaaS17 (BEN3). Infrastructure as a Service (IaaS) helps the organization to respond more quickly to change.
IaaS18 (BEN4). Infrastructure as a Service (IaaS) helps the organization provide better services to its clients.

Attitude - Infrastructure as a Service (IaaS) [27]
IaaS19 (ATT1). Using Infrastructure as a Service (IaaS) is a good idea.
IaaS20 (ATT2). Using Infrastructure as a Service (IaaS) is a wise idea.
IaaS21 (ATT3). I like the idea of using Infrastructure as a Service (IaaS).

Security – On Premise Service [23]
OnPrem1. I think that On Premise Service has mechanisms to ensure safe transmission of user information.
OnPrem2. I think that On Premise Service has sufficient technical capacity to ensure data will not be intercepted by unauthorized parties.
OnPrem3. I think that On Premise Service has sufficient technical capacity to ensure data will not be modified by unauthorized parties.

Trust – On Premise Service [24]
OnPrem4. On Premise Service is trustworthy.
OnPrem5. On Premise Service keeps promises and commitments.
OnPrem6. I trust On Premise Service because it has our best interests in mind.

Risk – On Premise Service [25]
OnPrem7. It is probable that On Premise Service would be worth its cost.
OnPrem8. It is probable that On Premise Service would frustrate me because of its poor performance.
OnPrem10. Comparing with other technologies, using On Premise Service has more uncertainties.

System Quality – On Premise Service [26]
OnPrem11. On Premise Service is easy to use.
OnPrem12. On Premise Service is user friendly.
OnPrem13. On Premise Service provides high availability.
OnPrem14. On Premise Service provides high-speed information access.

Perceived Benefits – On Premise Service [26]
OnPrem15. On Premise Service helps the organization save costs.
OnPrem16. On Premise Service helps the organization to increase return on investment.
OnPrem17. On Premise Service helps the organization to respond more quickly to change.
OnPrem18. On Premise Service helps the organization provide better services to its clients.

Attitude – On Premise Service [27]
OnPrem19. Using On Premise Service is a good idea.
OnPrem20. Using On Premise Service is a wise idea.
OnPrem21. I like the idea of using On Premise Service.

*Two control questions (IaaS9 and OnPrem9) are removed.

Understanding the Impact of Contextual Clues in Misinformation Detection

Kevin Matthe Caramancion
College of Emergency Preparedness, Homeland Security, and Cybersecurity
University at Albany, The State University of New York
Albany, New York, United States
e-mail: kcaramancion@albany.edu

***Abstract*—To further reduce the pervasiveness of fake news on social networking sites and increase content integrity, an important yet one of the most basic of defenses has be explored—content descriptors. This paper investigates the effects of content descriptors in aiding the human factor against disinformation attacks through its recognition. An explicit fake news test was employed to assess the users on how they use these descriptors as contextual clues in categorizing specific contents as either legitimate or misleading. The results of this study reveal the varying effect of contextual clues in a population to recognize fake news.**

Keywords-disinformation; misinformation; fake news; informatics; cybersecurity; threat

I. INTRODUCTION

A number of social networking sites such as Facebook and Twitter allow users and entities alike to create and/or share contents in their respective spaces. A particular type of content in these sites is news and journalistic information. More than being able to share contents in these sites, information users consume such crucial information from these media and as such the integrity of contents is of crucial importance.

On the other side of the spectrum, misleading information had been long existing in the same environment, persisting and mixed with the legitimate news with the intent to deceive users due to several varying reasons and inspirations, mostly political and conspiracy induced. Since it meets all the technical criteria to be tagged as such, [1] advocates for the categorization of fake news as a cybersecurity threat. Ultimately, the information consumer has the personal responsibility to decide for himself which of these are factual and which are misleading [2]. Thus, highlighting the ability to detect misinformation with the existing information descriptors, i.e. the contextual clues, is of extreme importance.

With social media feeds consisting of virtually unlimited information, how does the user filter and decide which is legitimate? The underlying main research question of this paper is How do contextual clues in news headlines affect users in detecting misinformation? Specifically, this paper will attempt to investigate the impact of contextual clues to users in news consumption. And if so, to what degree of an impact will it have: either a positive or negative effect to the detection of misinformation in the user's aid.

The literary significance of this study is through its contribution to the existing research and knowledge body in the domain of fake news and its effects on information users and their perception. The creation of a novel data instrument that empirically extracts the capability of users to categorize a content from a legitimate and misleading has been introduced. Furthermore, it adds to the current literary body through the highlighting of an underemphasized yet crucial aid in the proper categorization of a legitimacy of a content—contextual clues.

The pragmatic significance and outcome of this study is the assessment of the ability of individuals and groups to detect misinformation when presented with one. The precursor to the correction and removal of misleading contents in social spaces first relies on its recognition. Technologies that aim to automate the detection of fake news on social networking websites require trainings from humans. The transfer of detection skills from a user to a computer highlights the importance of accurate human judgement of these contents. An erroneous training data created and injected by humans to machines will later on result in an automated yet erroneous categorization of contents and vice versa.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

This paper will (1) discuss the literature highlighting the prevalence of disinformation, combatting fake news, and the role of contextual clues in human to machine transfer (2) discuss the methodology used to obtain and analyze data, including its control and limitations, (3) presentation of findings, (4) discussion and analysis of findings, and (5) highlight the prospectus for future research.

II. Literature Review

A. *Pervasiveness and Dangers of Disinformation*

The increasing damages produced by fake news have been documented on a global scale [3]. These come in the form of moral and financial damages along with the increase in polarization and disputes in online societies [4]. The increase in active daily users of social networking sites enabled and further amplified the resonating effects of disinformation attacks [5][6]. One particular cause to this is revealed by the findings of [7] and [8]'s study wherein users displayed over-reliance on social media sites not just for the purpose of entertainment but also for news consumption. These researchers highlight the drastic increase of user dependency on social media for all type of information.

On top of this, weaponized misinformation, i.e. disinformation, has been increasingly and consistently being used as a medium for political and conspiracy-charged propaganda by several actors and entities [9]. The most prevailing form of disinformation campaigns comes in the form of news and journalistic information. [10] further explained that most false news has hooks to captivate the readers with highly catchy headlines. These headlines are often lingering around social media websites including Facebook and Twitter [11][12]. Findings from [13] projects that damages from these attacks are just in premature phase and is estimated to exponentially increase should it not be managed, controlled and accounted for.

B. *Combatting Fake News*

Based from the Information Technology Infrastructure Library (2019)'s model, the series of responses different organizations and fields to disinformation attacks comes down into strengthening three distinct factors—people, process, and technology.

For the factor of people, awareness campaigns such as [14]'s are the suggestion mitigation response. It includes cybersecurity education that guides and equip users to be keen in the news and journalistic content they consume on social networking sites. Laws and policies have been introduced by organizations and national departments of different countries to support these [3][15]. The strategies instilled in the users include checking the credibility of source, date posted, and multimedia inscribed to the content as a support for its contents.

The factor of process includes the dynamics surrounding social media rules, strategies, and policies that regulate the integrity of content posted by users in social media websites [16][17]. These include total prohibition of advertisements that may mislead and outright removal of user accounts that produces such content.

Finally, technological factors involve the creation of computing innovations that detects misleading contents and have them categorized, corrected and at times, removal [18][19]. It is crucial to note that the underlying specialty fields responsible for these technological remedies are Artificial Intelligence and Machine Learning which both require training datasets for machines—all sourced from the collective experiences of users in detecting misinformation.

C. *Contextual Clues and Literary Gaps*

There is very little to almost virtually none existing study or research, however, that assesses the degree to which these descriptors impact an information user in performing this undertaking. Furthermore, no exhaustive list in this domain that explicitly establishes the types and forms of content descriptors which may assist an information user in this undertaking exists. Clearly, there is a need to further enrich this particular part of the domain since it is ideally supposed to function as an aid in combatting fake news. By taking into account the contextual clues' weakest link and strengthening its positively useful parts—it can be a potent form of defense against the ever-increasing misinformation

campaigns and attacks persisting in social media environments.

III. METHODOLOGY

Figure 1 below describes an overview of the study's methodology. Note that the novel data collection device is the fake news test.

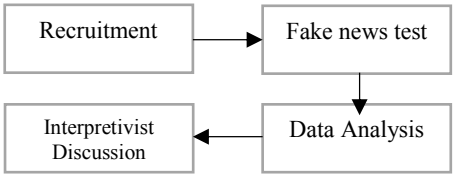


Fig. 1. Block Diagram of Methodology

A. The Population

The survey was opened for an undergraduate class of Introductory in Informatics. Of the 200 eligible and invited students to take the optional survey, 158 responded. Declared majors and concentrations of subjects are varying including but not limited to cybersecurity, computer science, informatics, economics, communications, sociology, business administration, accounting, and physics. Age range of subjects are from 18 to 25. The main instrument of this study is a mixture of Collins and Frenkel (2018)'s & Chua and Banerjee (2018)'s [20][21]. Collins and Frenkel (2018) published a web article in New York Times that displays a test intended for the viewers. Their simple quiz involves having to choose between two images without descriptors and from there—the viewer decides which is legitimate and which is misleading. Chua and Banerjee (2018)'s employed a slightly different approach in form of a survey. Their survey involves medical rumors (i.e. fake news)—deployed to medical professionals. Their survey involves a 5-point Likert scale that asks how likely the subjects will share the rumor to his/her social circle. Interestingly, the rumors in the survey are mixed with factual information and randomly assigned a "counter-rumor" as descriptors (i.e. contextual clues).

B. Data Collection

The main instrument of this study is a survey that explicitly asks a subject to categorize a content based on its legitimacy. Unlike Chua and Banerjee (2018)'s, this study's distinct feature is its more objective choices based from the former's 5-point Likert scale as a measurement. The answer choices for this study are a.) Legitimate (Highly Likely), b.) Misleading (Highly Unlikely), or c.) Unable to Detect (Somewhat in Between). Furthermore, to test the impact of the descriptors (i.e. contextual clues) on the subjects—the questions were displayed twice: One with context and one without context on all subjects. To further restrict bias, no control group is implemented. This is to capture the impact of contextual clues (as control variable) on the same content, on all respondents. The contextual clues included as descriptors in the contents are as follows: the source, the date it was added, and the external web link to the content as reference.

C. Data Analysis

General descriptive statistics and grade distribution were used to establish the baseline of measures of the results. Next, a comparison matrix is used to display the resulting differences of two same question content: One with contextual clues, and the other without contextual clues. A summary table was used display the principal findings of this study, the impact of contextual clues to the subjects. Finally, an interpretivist approach was employed by the author to give possible explanations of the results.

D. Controls and Limitations

The survey is deployed in blackboard and is strictly timed so as to mimic the actual stimuli environment of browsing a social media feed. The whole survey is timed 3 minutes and will auto-submit after the timer has elapsed. The range of user attention span per content as revealed by Facebook Research are varying from 1.7 seconds to 12 seconds. To take into account any unforeseen overhead on the subjects—an additional 60 seconds was added to the total time. The questions are sequenced in such a way that displays all the question contents without context clues first followed by the set with context clues. The themes of the questions are all relevant to the national and international affairs within the United States. Backtracking was prohibited to disallow any modification of answers to earlier questions after having viewed the questions with contextual clues

IV. RESULTS

Overview

Table 1 displays the general descriptive statistics of the results in aggregate form. Table 2 presents a grade distribution summary of the population mapped against the score—grouped in increments of 10s. Table 3 and Table 4 highlight the effects of contextual clues in each content mapped against each item of the test.

TABLE I. GENERAL STATISTICS

Respondent Count	158
Maximum Possible	10
Lowest Score	0
Highest Score	9
Range	9
Average	4.92405
Median	5
Standard Deviation	1.84371
Variance	3.39929

TABLE II. GRADE DISTRIBUTION

Greater than 100	0
90 - 100	1
80 - 89	9
70 - 79	21
60 - 69	33
50 - 59	33
40 - 49	33
30 - 39	12
20 - 29	7
10 - 19	5
0 - 9	4
Less than 0	0

TABLE III. COMPARISON MATRIX

WITHOUT CONTEXTUAL CLUES			
	LEGITIMATE	MISLEADING	UNABLE TO DETECT
ITEM1	13.92%	72.78%	10.13%
ITEM2	39.87%	38.61%	18.90%
ITEM3	56.96%	27.22%	12.66%
ITEM4	32.91%	36.08%	27.22%
ITEM5	51.27%	29.75%	15.19%

WITHOUT CONTEXTUAL CLUES			
	UNANSWERED	% CORRECT	% INCORRECT
ITEM1	3.16%	72.78%	27.22%
ITEM2	2.53%	39.87%	60.13%
ITEM3	3.16%	56.96%	43.04%
ITEM4	3.80%	32.91%	67.09%
ITEM5	3.80%	29.75%	70.25%

WITH CONTEXTUAL CLUES			
	LEGITIMATE	MISLEADING	UNABLE TO DETECT
ITEM1	44.30%	36.71%	12.03%
ITEM2	69.62%	13.29%	8.23%
ITEM3	74.05%	12.03%	3.16%
ITEM4	49.37%	22.15%	14.56%
ITEM5	34.12%	30.38%	16.46%

WITH CONTEXTUAL CLUES			
	UNANSWERED	% CORRECT	% INCORRECT
ITEM1	6.96%	36.71%	63.29%
ITEM2	8.86%	69.62%	30.38%
ITEM3	10.76%	74.05%	25.95%
ITEM4	13.93%	49.37%	50.63%
ITEM5	19.05%	30.38%	69.62%

TABLE IV. SUMMARY OF RESUTS AND DEGREE OF IMPACT

	CHANGE PERCENT	IMPACT	% INCORRECT
ITEM1	-36.08%	NEGATIVE	63.29%
ITEM2	29.75%	POSITIVE	30.38%
ITEM3	17.09%	POSITIVE	25.95%
ITEM4	16.46%	POSITIVE	50.63%
ITEM5	0.63%	POSITIVE	69.62%

V. DISCUSSION

A. *Discussion 1: Contextual clues have different effects on an information user's perception in detecting fake news*

Application of contextual clues to news headline will not necessarily have a positive effect on the user's part in recognizing misleading

contents. This is made visible by one resulting case wherein contextual clues further misled users and significantly reduced the accuracy of users in the detection.

B. *Discussion 2: Fake news backed with an unfamiliar reference—will make it appear legitimate*

Unfamiliarity of the source and its background further instilled doubts and misled users in the detection of fake news. This is made visible by one resulting case of a satire which was initially correctly tagged as misleading and later on garnered incorrect tags after being presented with contextual clues.

C. *Discussion 3: Timely and relevant fake news when supported by false photo and/or video appears to be very highly misleading*

Significant amount of incorrectly tagged contents (at most 70% error) were sourced from local and international news with relevant but false photo. As a matter of fact, after presenting these contents with contextual clues to the users—very little positive change occurred (less than 1% positive change).

D. *Discussion 4: Entertainment news, although legitimate, appears to be commonly perceived as misleading*

News surrounding public figure in the entertainment industry are often and generally assumed to be misleading when presented without contextual clues. Little change in this perception has shown to affect the users in the categorization of such content even after being presented with credible clues.

E. *Discussion 5: When a source is recognized and credible—it will usually clear the doubts and misconception of a content in question*

Contextual clues generally increased the accuracy of detecting misleading contents. The degree, however, of its positive impact is varying and the change resulted from the clues are of little (0.63%) to moderate significance (29.75%).

VI. Conclusion And Future Works

Mark Twain in his own words recounted that "*It is easier to fool people than to convince them that they have been fooled*." This study investigated the role and impacts of contextual clues in aiding a user in detecting misinformation. Like a double-edge sword, contextual clues can enlighten or further astray users in accurately categorizing the legitimacy of a content. They are created with the purpose similar to metadata—to describe the principal data. Usually it results in positive influence on users however it also has the potential to significantly mislead when users are not familiar with properly consuming it—the paradox revealed by this study. The burden of validation still lies on the user to responsibly examine the integrity of every content on a social environment.

Future research may examine the correlation of the demographics and socioeconomic factors of the users mapped against the results of the novel data instrument introduced by this study. The study is cross-sectional, so it only captures a population's assessment at one point in time due to the fast-paced nature of journalistic events and news information. Future research may consider adapting longitudinal approach as instances of shifts in knowledge of global events and affairs may affect the resulting perception of the users in recognizing misinformation. Another key consideration for future studies may be the shift of perspective from the users to the content creators on how to appropriately design contents and appropriately supplementing it with contextual clues to reduce misinformation.

Acknowledgment

This paper is especially dedicated to my grandfather and nana: Both of whom are always falling victimized by fake news. I promise I will help create a world of truth and one that is less dark—for you and all our people.

And most importantly to my other half, Mr. Vieira, first for being my inspiration and for being my patient zero in testing the contents of my fake news test.

References

[1] Caramancion, K. M. (2020, March). An Exploration of Disinformation as a Cybersecurity Threat. In *2020 3rd International*

Conference on Information and Computer Technologies (ICICT) (pp. 440-444). IEEE.

[2] Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, *57*(2), 102025.

[3] Klein, D., & Wueller, J. (2017). Fake news: a legal perspective. *Journal of Internet Law (Apr. 2017)*.

[4] Rainie, H., Anderson, J. Q., & Albright, J. (2017). The future of free speech, trolls, anonymity and fake news online.

[5] Chaffey, D. (2020). Global Social Media Research Summary. *Smart Insights*. Accessed via smartinsights.com

[6] Barberá, P., Tucker, J. A., Guess, A., Vaccari, C., Siegel, A., Sanovich, S., ... & Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature.

[7] Hermida, A., Fletcher, F., Korell, D., & Logan, D. (2012). Share, like, recommend: Decoding the social media news consumer. *Journalism studies*, *13*(5-6), 815-824.

[8] Pentina, I., & Tarafdar, M. (2014). From "information" to "knowing": Exploring the role of social media in contemporary news consumption. *Computers in Human Behavior*, *35*, 211-223.

[9] Marwick, A., & Lewis, R. (2017). Media manipulation and disinformation online. *New York: Data & Society Research Institute*.

[10] Benkler, Y., Faris, R., & Roberts, H. (2018). Network propaganda: Manipulation, disinformation, and radicalization in American politics. Oxford University Press.

[11] Marchi, R. (2012). With Facebook, blogs, and fake news, teens reject journalistic "objectivity". *Journal of Communication Inquiry*, *36*(3), 246-262.

[12] Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science*, *363*(6425), 374-378.

[13] Roguski, P. (2019). Overcoming the damage of disinformation. *New Eastern Europe*, *35*(1), 69-76.

[14] Torres, R., Gerhart, N., & Negahban, A. (2018, January). Combating fake news: An investigation of information verification behaviors on social networking sites. In *Proceedings of the 51st Hawaii International Conference on System Sciences*.

[15] Toronto, W. D. (2018). Fake News and Kill-Switches: The US Government's Fight to Respond to and Prevent Fake News. *AFL Rev.*, *79*, 167.

[16] Mosseri, A. (2017). Working to Stop Misinformation and False News. *Facebook for Media*. Accessed via facebook.com/facebookmedia

[17] Shapiro, A. (2020). Twitter Vows That as Disinformation Tactics Change, Its Policies Will Keep Pace. *NPR*. Accessed via npr.org

[18] Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, *52*(1), 1-4.

[19] Ahmed, H., Traore, I., & Saad, S. (2017, October). Detection of online fake news using N-gram analysis and machine learning techniques. In *International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments* (pp. 127-138). Springer, Cham.

[20] Collins, K. & Frenkel, S. (Sept. 4, 2018). Can you spot the deceptive Facebook post? *New York Times*. Accessed via nytimes.com

[21] Chua, A. Y., & Banerjee, S. (2018). Intentions to trust and share online health rumors: An experiment with medical professionals. *Computers in Human Behavior*, 87, 1-9.

A Generative Modelling Technique for 3D Reconstruction from a Single 2D Image

Saurabh Kumar Singh
Dept. of Computer Science & Engineering
IIT (BHU), Varanasi
Varanasi, India
saurabhk.singh.cse16@iitbhu.ac.in

Shrey Tanna
Dept. of Computer Science & Engineering
IIT (BHU), Varanasi
Varanasi, India
shrey.tanna.cse16@iitbhu.ac.in

***Abstract*—3D Object Reconstruction is the task of predicting the 3D model of an object given a set of 2D images. In this paper, we propose an approach to solving this problem, given a single 2D image. We attempt to make use of several deep learning techniques. Our model consists of two parts. The first part generates multiple images having different viewpoints. We have included this part because reconstructing 3D object directly from a single 2D image is quite difficult, but the same task would be a lot easier given multiple images which capture different views of that same object. Also, predicting an image having a different viewpoint is much easier than predicting the whole 3D object, given an input image. The second part uses a network consisting of an Encoder, a Decoder (or Generator), and a Discriminator to predict the complete 3D voxel grid of the object. In this way, we achieve significant improvements in the results as compared to the existing techniques.**

***Index Terms*—Reconstruction, GANs, CNNs, Neural networks, Voxel**

I. Introduction

The rapid development of fields like robotics, designing, virtual reality, medical imaging, etc. requires a crucial understanding of shapes of objects, and hence 3D understanding of objects has itself evolved as an interesting problem for computer vision researchers. As the computer storage space is increasing, the visual data is also increasing proportionally, and hence studying their shape features has been a worry. The problem would have been easier if somehow we could extract a 3D model out of it. This creation of a 3D model from a given set of image(s) is called 3D reconstruction. Let us understand this with an example. Say, there is a bus standing on the road, and we click a picture of it using a camera. The picture we get is a 2D image. If we are given this photo and asked to predict its 3D model, it is quite tough. It can be seen that 3D reconstruction is just the reverse of clicking a 2D photo of a 3D scene. When we click a picture of an object using a camera, we only get its projection onto a given plane. The loss of most of the depth information makes the problem challenging. A human, when asked to predict the corresponding 3D counterpart of a given image, as he/she may have some idea about that as he/she has seen several similar pictures throughout his/her life. This gives us a hint that the machine can be trained with lots of visual data to accomplish the task. As we know that a given point on a 2D image may have any depth on the 3D model, it is very difficult to predict the depth using only one image. However, the position of the point can be found as the intersection of the projection rays provided we have two images of that object.

In this paper, we present a two-stage approach. The first one focuses on generating several images having different viewpoints from a single, using convolutional neural networks. The second one constructs a 3D model from these multiple view images using generative adversarial networks. This paper presents an improved model of present deep-learning methods [1] [2], which is well reflected in our results.

The organization of the paper is as follows: Section II presents a summary of the literature survey showing different methods for multiple-view image generation and non-neural network based and neural network based methods of 3D reconstruction. The proposed approach is discussed in Section III. The experimental analysis and results are discussed in Section IV. Finally, the paper is concluded in Section V.

II. Related Work

A. Multiple views prediction from a single image

Several approaches have been studied for obtaining unseen views of a given image using different image transformation or neural network methods. Transforming autoencoders [3] show us how neural networks deal with variations in orientation, scale, position and lighting of the image. The Deep Convolution Inverse Graphics Network (DCIGN) [4] can generate images with a different pose for the same object and similarly, images with different lighting given an input image. It consists of multiple layers of convolution and de-convolution operators and is trained using the Stochastic Gradient Variational Bayes (SGVB) algorithm [5]. Such task is much popular in the field of face recognition and manipulation, as various factors like identity, view and illumination are coupled in the face images and hence disentangling such factors is a major challenge. As an example, the Multi-view perceptron (MVP) [6] disentangles the face identity and predict features under different viewpoints. Also, another similar problem is comparing two images from different views, which turns out to be quite challenging as features are not stable under large view point changes. [7] focuses on solving this problem by synthesizing the features for other views with the help of 3D model collection of related

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

objects, based on which feature sets for the two images is created and compared. To render rotated objects from a single image, [8] is an end-to-end trained network. Given an image, it basically gives us the view from a viewpoint differing with a fixed angle. As a result of this, a discrete set of views of the given image may be generated by using the algorithm again and again. However, it may lead to error accumulation in case of large angle. As opposite of this, [2] can generate views, varying the angle continuously.

B. Neural Network Based 3D Reconstruction

1) RNN-based 3D object reconstruction: Recurrent neural networks have been used for reconstructing the 3D object from both single and multi-images input. The paper [9] proposes a novel recurrent neural network architecture, called 3D Recurrent Reconstruction Neural Network (3D-R2N2) for 3D object reconstruction. In this network, we can feed one image or multiple images. These images may be taken from randomly oriented viewpoints. In this whole procedure, no pre-processing elements like segmentation, viewpoint labels, or key points are needed. This method uses recurrent neural networks like LSTMs and GRUs, and convolutional neural networks as well.

The network aims at performing object reconstruction from single or multi-view images. The network consists of three components, a 2D Convolutional Neural Network (2D-CNN), a 3D Convolutional LSTM (3D-LSTM), and a 3D Deconvolutional Neural Network (3D-DCNN).

2) CNN-based 3D object reconstruction: Convolutional neural networks are used in most of the computer vision problems. Here, we are discussing some of the methods using only CNNs for 3D reconstruction. The paper [10] explores the point set representation of a 3D object using a single image of that object. A point cloud representation as output is chosen as it is a simple and uniform representation and allows geometric transformations or deformations easily with some manipulations. Also, in cases where ground truth may be ambiguous, this representation provides us the best output. To be capable of this, the model analyzes the visible parts of the object with the help of the input image and then cleverly guesses the rest part.

The network aims at outputting a point set representation of a 3D object given a single image of that object as input. 3D shape of the object is represented as an unordered point set $S = (x_i; y_i; z_i)_{i=1}^{N}$ where N is a constant. Mostly, N is taken as 1024. Such representation can have geometric transformations (like rotation, scaling, translation or their combinations) easily by simple matrix algebra. There is an encoder in the network and a decoder(or a predictor) in the network.

3) GAN based 3D Reconstruction: Generative adversarial networks [11] have several advantages over the other methods and can produce high-quality realistic objects and outperforms several other methods in terms of performance. There are other modified versions of GANs available, as mentioned further. WGAN or IWGAN [12] mentions a method for stable training of GANs by providing an alternative training method, which includes weight clipping or enforcing Lipschitz constraint. DCGAN [13] uses convolutional networks for implementing GANs. Stacked GAN [14] is trained to invert the hierarchical representations of a bottom-up discriminative network and is able to generate images of much higher quality than GANs without stacking. Different methods for improvement of GANs are also described in [15], [16] and [17]. [1] mentions an approach generating a 3D object from a probabilistic space using GANs and CNNs. [18] also achieves this, but uses IWGAN for the same. MarrNet [19] uses 2.5D sketches such as a normal map, depth map, and silhouette map in addition to single image are input to the network in order to improve 3D-GAN [1]. Most of this 3D work uses the Shapenet dataset [20].

[21] trained generative up-convolutional neural networks which are able to generate images of objects given object style, viewpoint, and color. [22] presented a method for joint analysis and synthesis of geometrically diverse 3D shape families. [23], [24] proposed to learn a joint embedding of 3D shapes and synthesized images, [25], [26] focused on learning discriminative representations for 3D object recognition, [27], [28], [29] discussed 3D object reconstruction from in-the-wild images, possibly with a recurrent network, and [24], [30] explored autoencoder-based networks for learning voxel-based object representations. [8] proposed a novel recurrent convolutional encoder-decoder network that is trained end-to-end on the task of rendering rotated objects starting from a single image.

C. Non-neural Network Based 3D Reconstruction

A lot of approaches are also available, which does not involve any use of neural network techniques.

[31] presented an automated pipeline with pixels as information sources and 3D surfaces of different classifications as a result of pictures of some scenes. Their methodology had deformable 3D models that can be learned from 2D annotations available in existing object detection data sets. [32] proposed a two-step method. Initially, they employ orthogonal matching pursuit to choose the closest single CAD model in the dictionary to the projected image. Finally, they use their graph embedding based on local dense correspondence to allow for sparse linear combinations of the CAD models.

III. Proposed Approach

In this section, we introduce our approach to solving the 3D Reconstruction problem. We can easily divide our approach into two parts which are discussed in the following sections.

A. Multiple view images generation step

The first part of the approach is to generate multiple view images for the given image. Several images having different viewpoints are produced using the multi-view 3D (mv3D) CNN network inspired by the paper [2]. This network takes an image as well as a viewpoint vector as input and outputs the required image. Viewpoint is described using a viewpoint vector containing five elements, two of which being the sine

and cosine of the azimuthal angle, the other two being the sine and cosine of the elevation angle, and the fifth one being the distance from the object center.

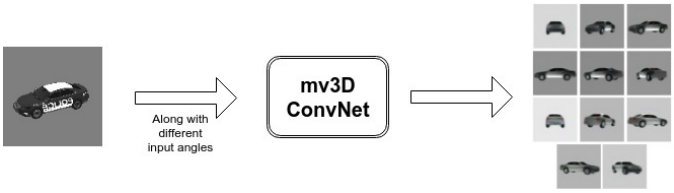


Fig. 1. First part of the network, showing multi-view images as the output of the mv3D network, given a single input image along with different viewpoint vectors

The network used is an encoder-decoder network. A pair of image and viewpoint vector (x_i, v_i) is input to the network, and the network aims to predict the required image (ground truth = f_i). The encoder consists of five convolutional layers (stride = 2) followed by a fully connected layer at the end. The viewpoint vector is independently processed by fully connected layers, and the result is merged with the encoder output. The decoder then takes the resultant vector through five deconvolutional layers. Deconvolution is the reverse of convolution. Each pixel is replaced with a 2×2 pixel window with original pixel value in the top-left pixel and filling others with zero.

To train the network, any pair of snapshots from the 3D model of some object along with angle information is taken as an input-output pair. The training is done by minimizing the loss function, which is squared Euclidean loss for the required image, and as shown below:

$$L = \sum_i \parallel f_i - \hat{f}_i \parallel_2^2 \tag{1}$$

where $\hat{f}_i$ is the output image and f_i is the ground truth image for input image x_i and viewpoint vector v_i.

For a given image x_i, we make 11 pairs of inputs as (x_i, v_{i1}), (x_i, v_{i2}), .., (x_i, v_{i11}) using 11 different viewpoint vectors v_{i1}, v_{i2}, .., v_{i11} respectively. It can be seen in the Fig. 1 that 11 images of different viewpoints are produced in our implementation, however, this number can be changed appropriately. Now, the original image x_i, along with 11 images $\hat{f_{i1}}, \hat{f_{i2}}, .., \hat{f_{i11}}$ produced, is taken as a whole to form a $256 \times 256 \times 12$ block by taking only the L-channel of the respective images. This block is then taken as the input for a 3D-GAN network, which constitutes the second part of our solution as described in the section as under.

B. 3D model construction step

In this section, we discuss the second part of the network, which takes in the output of the first part of the network as input, and outputs the required 3D voxel grid.

The network consists of three components: Encoder (E), Generator (G), and Discriminator (D). The encoder takes the $12 \times 256 \times 256$ block (a set of 12 images' L-channels) as input. It consists of five 3D convolutional layers with number of channels, kernel size, and strides respectively, as $(64, 11, 4)$, $(128, 5, 2)$, $(256, 5, 2)$, $(512, 5, 2)$, $(400, 8, 1)$. Finally, we sample the output of the last layer (400-sized vector) into two sets to form a mean vector and a standard deviation vector, giving rise to a 200-dimensional vector. This vector is further used as input to the generator. The generator (or decoder) consists of three 3D convolutional layers with number of channels, kernel size, and strides respectively, as $(128, 10, 2)$, $(64, 4, 2)$, $(1, 4, 2)$. It outputs a $20 \times 20 \times 20$ voxel grid, which is taken as input by the discriminator to check whether it is real or fake by estimating a score using ground truth. The discriminator consists of three 3D convolutional layers with number of channels, kernel size, and strides respectively, as $(64, 4, 2)$, $(128, 4, 2)$, $(2, 2, 1)$. It outputs a value between 0 and 1, denoting the probability of the 3D voxel being real/fake. We also add batch normalization layer and the layer having leaky ReLU units with parameter 0.2 in between convolutional layers throughout.

The overall loss function is made up of the following components.

The first one is the 3D-GAN loss function, which is a minimax loss as it involves the update of two components G and D, taking turns alternatively by minimizing generator loss and maximizing discriminator loss. The Discriminator aims to maximize the probability of assigning the correct label to the training examples and output of G. So, we can write the loss function for D as:

$$L_{Disc} = -E_x[\log(D(x))] - E_z[\log(1 - D(G(z)))] \tag{2}$$

where E is the mathematical expectation operator, x is 3D voxel grid input to D, z is the vector (output of E) input to G. The generator aims to minimize the inverse probability by D for fake samples. So, we can write the loss function for G as:

$$L_{Gen} = E_z[\log(1 - D(G(z)))] \tag{3}$$

The second one is the Kullback-Leibler divergence loss, which aims to confine the distribution space of output of the encoder E.

$$L_{KL} = D_{KL}(N(\mu, \sigma) \parallel N(0, I)) \tag{4}$$

where $D_{KL}(P \parallel Q)$ is the divergence between P and Q, $N(\mu, \sigma)$ denotes the normal distribution with mean μ and standard deviation σ (produced by the encoder), and $N(0, I)$ denotes the standard normal distribution.

The third loss function represents the 3D model reconstruction loss, which is the squared euclidean distance between the generator output and the target shape. It is given by:

$$L_{3D} = \parallel y - \hat{y} \parallel_2^2 \tag{5}$$

where y is the target 3D shape of an image, and $\hat{y}$ is the 3D output of the generator.

In each iteration of the training step, an appropriate set of loss functions are used for updating a particular component of the network. We have used ADAM optimizer for training purpose, setting the parameters as $\alpha = 0.00008$, $\beta_1 = 0.5$

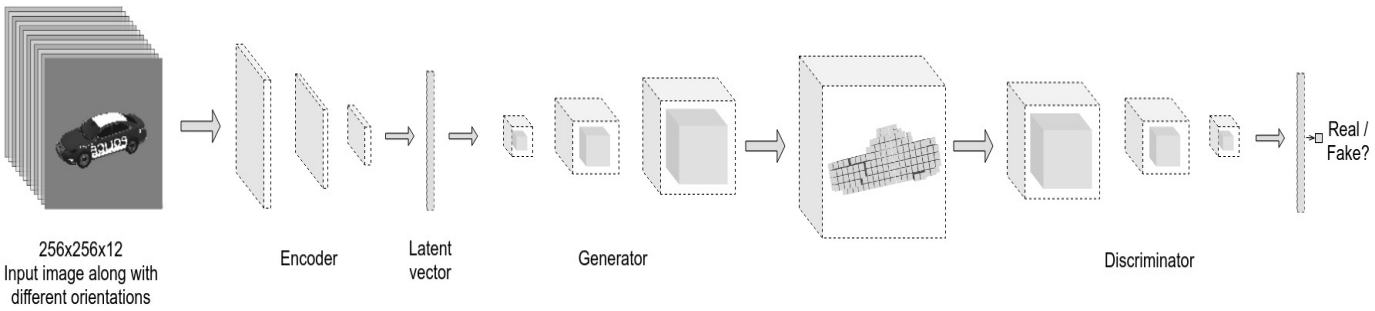


Fig. 2. Second part of the network, showing its three components: Encoder, Generator and Discriminator

and $\beta_2 = 0.9$. We have trained the algorithm for 1200 epochs, keeping the batch size to be 256.

IV. Experiments

A. Dataset and System Description

The 3D CAD models are downloaded from the Shapenets [20] website. These models are then rendered as images, given different illumination levels, size, azimuthal, and elevation angles. The images are then resized by cropping them from the center to make it 256×256 sized. These images and their corresponding voxel grid are used as the dataset for the task. We have trained the model for the car dataset.

We have used a system with three Graphics Processing Units (GPUs) for training purposes. One of them is Nvidia Titan Xp with 12GB RAM, 12196 MB FB memory and 256 MB BAR1 memory, and the other two are Nvidia GeForce GTX 1080 Ti with 11 GB RAM, 11178 MB FB memory and 256 MB BAR1 memory.

B. Results

We show the qualitative results for 3D reconstruction from a single image in fig. 3. In each row, the image on the left is the input image, and the other images show two arbitrary views of the corresponding 3D output. The results as can be seen, never miss a major part of the article even if it's thin, which was not the case with some other methods. This can easily be seen in the output for the tractor.

V. Conclusion and Future Work

In this paper, we present an approach that simplifies the problem of reconstructing a 3D object from a single image, to a problem of reconstructing a 3D object from multiple images(captured from different angles). For this task, we use the multi-view 3D CNN network, inspired by [2]. And the actual 3D generation after this step is done by the GAN. The encoder of this GAN takes the L-channels of several multi-view images as it's input. The decoder consists of three 3D convolutional layers, which outputs a voxel grid.

The approaches proposed till now and even our approach are likely to work for a single class image. The network is trained for a given class of objects and gives the desired output only in case of the images taken from that class. So, a robust solution that could work for all classes of images is to be developed and is a challenging problem.

Acknowledgements

This paper and the project would not have been possible without our mentor's support, Dr. Pratik Chattopadhyay, Assistant Professor, CSE, IIT(BHU), Varanasi. His valuable discussions, insightful suggestions, and knowledge kept our work on track.

References

[1] Jiajun Wu, Chengkai Zhang, Tianfan Xue, Bill Freeman, and Josh Tenenbaum. Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling. In *Advances in neural information processing systems*, pages 82–90, 2016.

[2] Maxim Tatarchenko, Alexey Dosovitskiy, and Thomas Brox. Multi-view 3d models from single images with a convolutional network. In *European Conference on Computer Vision*, pages 322–337. Springer, 2016.

[3] Geoffrey E Hinton, Alex Krizhevsky, and Sida D Wang. Transforming auto-encoders. In *International Conference on Artificial Neural Networks*, pages 44–51. Springer, 2011.

[4] Tejas D Kulkarni, William F Whitney, Pushmeet Kohli, and Josh Tenenbaum. Deep convolutional inverse graphics network. In *Advances in neural information processing systems*, pages 2539–2547, 2015.

[5] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[6] Zhenyao Zhu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Multi-view perceptron: a deep model for learning face identity and view representations. In *Advances in Neural Information Processing Systems*, pages 217–225, 2014.

[7] Hao Su, Fan Wang, Li Yi, and Leonidas Guibas. 3d-assisted image feature synthesis for novel views of an object. *arXiv preprint arXiv:1412.0003*, 2014.

[8] Jimei Yang, Scott E Reed, Ming-Hsuan Yang, and Honglak Lee. Weakly-supervised disentangling with recurrent transformations for 3d view synthesis. In *Advances in Neural Information Processing Systems*, pages 1099–1107, 2015.

[9] Christopher B Choy, Danfei Xu, JunYoung Gwak, Kevin Chen, and Silvio Savarese. 3d-r2n2: A unified approach for single and multi-view 3d object reconstruction. In *European conference on computer vision*, pages 628–644. Springer, 2016.

[10] Haoqiang Fan, Hao Su, and Leonidas J Guibas. A point set generation network for 3d object reconstruction from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 605–613, 2017.

[11] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[12] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C Courville. Improved training of wasserstein gans. In *Advances in Neural Information Processing Systems*, pages 5767–5777, 2017.

[13] Wei Fang, Feihong Zhang, Victor S Sheng, and Yewen Ding. A method for improving cnn-based image recognition using dcgan. *Comput. Mater. Contin*, 57:167–178, 2018.

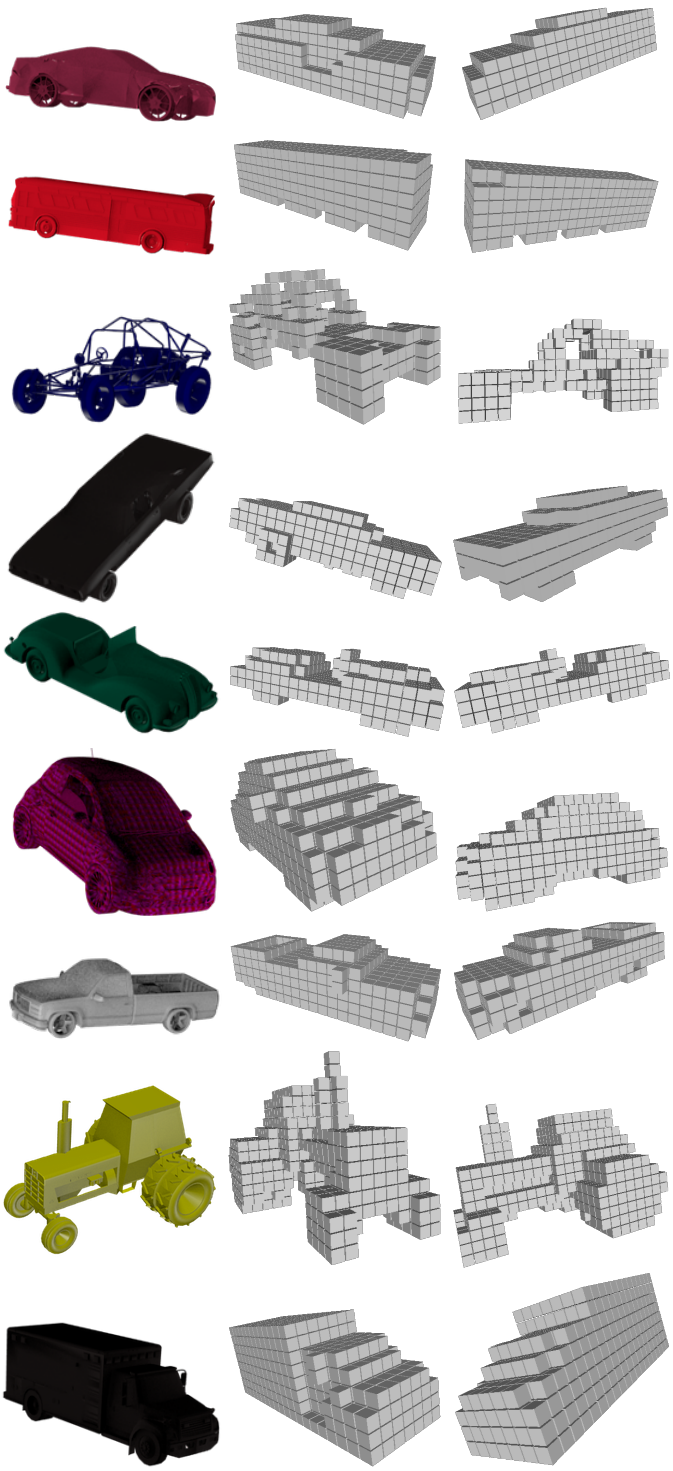


Fig. 3. Results showing 3D reconstruction from a single image. In each row, the image on the left is the input image, and the other images show two arbitrary views of the corresponding 3D output

[14] Xun Huang, Yixuan Li, Omid Poursaeed, John Hopcroft, and Serge Belongie. Stacked generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5077–5086, 2017.

[15] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation. *arXiv preprint arXiv:1710.10196*, 2017.

[16] Ian Goodfellow. Nips 2016 tutorial: Generative adversarial networks. *arXiv preprint arXiv:1701.00160*, 2016.

[17] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. In *Advances in neural information processing systems*, pages 2234–2242, 2016.

[18] Edward Smith and David Meger. Improved adversarial systems for 3d object generation and reconstruction. *arXiv preprint arXiv:1707.09557*, 2017.

[19] Jiajun Wu, Yifan Wang, Tianfan Xue, Xingyuan Sun, William T Freeman, and Joshua B Tenenbaum. MarrNet: 3D Shape Reconstruction via 2.5D Sketches. In *Advances In Neural Information Processing Systems*, 2017.

[20] Angel X Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, et al. Shapenet: An information-rich 3d model repository. *arXiv preprint arXiv:1512.03012*, 2015.

[21] Alexey Dosovitskiy, Jost Tobias Springenberg, Maxim Tatarchenko, and Thomas Brox. Learning to generate chairs, tables and cars with convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(4):692–705, 2017.

[22] Haibin Huang, Evangelos Kalogerakis, and Benjamin Marlin. Analysis and synthesis of 3d shape families via deep-learned generative models of surfaces. In *Computer Graphics Forum*, volume 34, pages 25–38. Wiley Online Library, 2015.

[23] Yangyan Li, Hao Su, Charles Ruizhongtai Qi, Noa Fish, Daniel Cohen-Or, and Leonidas J Guibas. Joint embeddings of shapes and images via cnn image purification. *ACM transactions on graphics (TOG)*, 34(6):234, 2015.

[24] Rohit Girdhar, David F Fouhey, Mikel Rodriguez, and Abhinav Gupta. Learning a predictable and generative vector representation for objects. In *European Conference on Computer Vision*, pages 484–499. Springer, 2016.

[25] Hao Su, Charles R Qi, Yangyan Li, and Leonidas J Guibas. Render for cnn: Viewpoint estimation in images using cnns trained with rendered 3d model views. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2686–2694, 2015.

[26] Charles R Qi, Hao Su, Matthias Nießner, Angela Dai, Mengyuan Yan, and Leonidas J Guibas. Volumetric and multi-view cnns for object classification on 3d data. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5648–5656, 2016.

[27] Jiajun Wu, Tianfan Xue, Joseph J Lim, Yuandong Tian, Joshua B Tenenbaum, Antonio Torralba, and William T Freeman. Single image 3d interpreter network. In *European Conference on Computer Vision*, pages 365–382. Springer, 2016.

[28] Yu Xiang, Wongun Choi, Yuanqing Lin, and Silvio Savarese. Data-driven 3d voxel patterns for object category recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1903–1911, 2015.

[29] Christopher B Choy, Danfei Xu, JunYoung Gwak, Kevin Chen, and Silvio Savarese. 3d-r2n2: A unified approach for single and multi-view 3d object reconstruction. In *European conference on computer vision*, pages 628–644. Springer, 2016.

[30] Abhishek Sharma, Oliver Grau, and Mario Fritz. Vconv-dae: Deep volumetric shape learning without object labels. In *European Conference on Computer Vision*, pages 236–250. Springer, 2016.

[31] Abhishek Kar, Shubham Tulsiani, Joao Carreira, and Jitendra Malik. Category-specific object reconstruction from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1966–1974, 2015.

[32] Chen Kong, Chen-Hsuan Lin, and Simon Lucey. Using locally corresponding cad models for dense 3d reconstructions from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4857–4865, 2017.

Interactive Visualization and App Development for Precipitation Data in Sub-Saharan Africa

Divydharshini Karthikeyan[1], Sudha Shah[1], Aparna S. Varde[1,3] and Clement Alo[2,3]
1. Department of Computer Science
2. Department of Earth and Environmental Science
3. Environmental Management and Science PhD program
(karthikeyand1 | shah37 | vardea | aloc) @ montclair.edu

***Abstract*—Precipitation plays a highly significant part in the hydrological cycle and is responsible for depositing most of the freshwater on earth. It can also impact the climate ecosystem. Considerable work has been performed to understand, model and predict precipitation. However, to the best of our knowledge, hardly any interactive visualizations exist that allow end-users to obtain a better perception of local impacts; also there are no specific apps for this purpose. The two goals of our work are thus: to design a visualization tool that illustrates changes in precipitation in a given time period with analysis; and to design a mobile application (app) that displays Sub-Saharan precipitation patterns, both using big data. This work would be advantageous to environmental scientists, as well as local residents of the respective areas. It would also be beneficial to students interested in conducting climate-based research. It addresses the theme of IoT (Internet of Things) which involves providing ubiquitous access to Internet-Based information on platforms such as smartphones and other mobile devices.**

Keywords—Android App, Big Data, Data Visualization, Human Computer Interaction, Internet of Things, Data Mining

I. INTRODUCTION

Background: Climate change can impact society by increasing health-related risks, causing lack of freshwater availability, raising sea levels, and contributing to economic impacts. A warming climate is expected to influence the amount of precipitation recorded across the globe. Climate change can also affect weather patterns and precipitation levels in different locations. This could mean that some areas may see more precipitation and get exposed to flooding, whereas others may experience less precipitation and become affected by drought [1]. Over the last 100 years, precipitation has increased by an average of about 1% on all the land surfaces globally. In general, too much precipitation can impose concerns such as flooding and landslides, resulting in property damage and loss of life. Too little can pose issues as well, leading to droughts, wildfires, poor air quality, agricultural damage, and health degradation, resulting in an overall increase in drought risk and severity. There are many reasons for changes in precipitation. The primary cause is a change in temperature. Many scientists believe that rising temperature could increase the rates of evaporation from soil and water, as well as transpiration from plants, leading to an increase in the amount of precipitation [2]. For climatic research, precipitation is an essential component for equalizing the energy budget, and one of the primary challenging traits of climate modeling.

Motivation: In this paper, we focus on precipitation data from the Sub-Saharan region in Africa, i.e. located south of the Sahara Desert, a region that is highly impacted by climate change with respect to precipitation. While there are many studies conducted on precipitation data [1, 2] along with the concerned climatic models, to the best of our knowledge there is no interactive visualization tool or mobile application (app) for the purpose of analyzing and disseminating Sub-Saharan precipitation data. It has been indicated by farmers in this region that lack of access to information is a hurdle in the path of adapting to climate change, especially in precipitation. We deal with big data on precipitation of the order of TB (Terabytes). Big data is a widespread paradigm today that renders ample data and possibilities to enhance research [3]. Though big data provides the opportunities for advancements in many fields such as business, education and healthcare, it has various challenges including storage, transmission and processing [4]. Based on this motivation pertaining to precipitation in Sub-Saharan Africa and the big data therein, we provide our problem definition as follows.

Problem Definition: The problem addressed in this paper relates to visualizing and analyzing precipitation data in Sub-Saharan Africa along with its ubiquitous dissemination in a user-friendly manner. The goals of our work are as follows.

- *Interactive visualization with analysis of big data on precipitation studies in Sub-Saharan Africa*
- *Dissemination of the precipitation data via mobile application (app) development for ubiquitous access*

We describe our solution approaches, with a brief survey of related work, processing of big data on precipitation, design of the precipitation visualization tool for public access, and the development of an Android app for precipitation studies based on HCI methods. We also discuss our ongoing work.

II. RELATED WORK

Precipitation studies have been conducted by various researchers. Contractor et al. [5] introduce a novel gauge-based dataset with gridded daily precipitation of global and temporal land coverage from 1950 to 2013 named REGEN, the Rainfall Estimates on a Gridded Network. According to REGEN, there are notable changes in SSA (Sub-Saharan Africa) rainfall features between 1950 and 2013 in well-gauged areas. Contractor et al. [5] also examine rainfall trends and discuss variations of the trend based on different region types including wet, intermediate and dry. A comparison of REGEN and other gridded and satellite products is also performed. Hoscilo et al. [6] determine the indicators of rainfall based on satellite data. They determine vegetation trends in Africa based on both climatic and non-climatic characteristics. This article is able to identify the regions in Sub-Saharan Africa that experience positive, negative, or stable rainfall over a period of 10 years from 2001 to 2010. Barrios et al. [7] examine the impact that climatic variation has on the pattern of urbanization in Sub-Saharan Africa. The dependency of Sub-Saharan African agriculture on the precipitation and its indications for Africa's economic advancement are broadly discussed. Juana et al. [8] provide a

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

discussion on the challenges that farmers' face due to drought in the Sub-Saharan region, and adaptation to address these challenges, e.g. planting tree crops. The researchers also propose that most farmers in Sub-Saharan Africa are conscious that the region is getting warmer, and that the rainfall patterns have changed. Farmers believe that the lack of access to information about climate change is the main impediment to climate adaptation. Muller [9] present that global warming concerns and climate-related variations are likely to increase weather-related hazards and difficulties facing human settlements such as floods. The primary focus of the paper is on inadequate water conditions in Sub-Saharan region. The hurdles encountered by these countries are economic, institutional, and technological. They have made matters much harder given their financial circumstances.

While these studies consider various aspects of data on precipitation, to the best of our knowledge none of them provide user-friendly interactive visualization tools, nor do they provide mobile applications for data dissemination with ubiquitous access. Our work that addresses these two goals thus makes a contribution to precipitation studies, as well as the fields of Visualization and IoT.

Some of the work in this paper relates to our earlier work on app development [10, 11] and related areas [12 – 17]. We have developed an app for local water management in the NJ Passaic region considering the parameters of temperature, humidity and precipitation [10]. Another app [11] entails disseminating information on mining ordinances (local laws) and their public reactions, focusing on the greater NYC area, especially with the goal of assessing how well the given region heads towards a smart city. This is based on discovering interesting knowledge from ordinances and the public tweets about them, conducted in our earlier research [12, 13]. Other relevant work deals with mining worldwide data from multiple cities to derive information useful for population relocation with the aim of sustainable living [14]. There is some research on discovering useful knowledge from GIS (geographic information system) data pertaining to urban sprawl [15], i.e. uncontrolled expansion of urban areas, along with its causes and relevant decision support. Some of our research pertains to data centers and their greenness [16, 17], analyzing parameters such as carbon footprint and power usage effectiveness, in order to build a DSS (decision support system) for catering to sustainability goals. Our work in this paper is orthogonal to such research, making contributions to precipitation and its impacts on sustainability via information visualization and dissemination.

III. Description of Methods

A. Approach for Data Visualization

Visualization improves the ability to both examine and understand data, thus it significantly aids in decision-making and problem-solving [18]. When users can interact with data and answer predefined questions, or queries that arise during analysis, big data visualizations can reveal their full strength to commence the exploration of new associations and trends for further use [19]. Yet, it can be difficult to visualize big data due to its volume, variety etc. It can be necessary to downscale the data to make the most out of the available information and utilize it to its full potential [20].

In order to develop an effective visualization tool, it is crucial to modify the features to address the target audience. Our visualization method aims to address skilled researchers as well as naïve users without technical prowess. The key features of the visualization tool are thus designed in a manner that allows every user to interact with it conveniently, understand it easily, and access it regularly.

The usefulness of the data visualization tool is seriously influenced by the manner in which it is executed. Thus, we compile visualizations into dashboards with controls and filters. Static visualizations are used here, they can be updated to accommodate dynamic visualizations later.

B. Approach for App Design

For mobile applications, user expectations play a vital role and their requirements should be satisfied in order to provide an enjoyable and prolific experience. Good interface design always attracts users, hence it is essential to make the application look attractive, be quickly accessible and include features that are easily comprehensible. This is where Human Computer Interaction (HCI) comes into play for designing and developing effective apps [21].

The initial step of the design process involves creating a sketch of the outline and is followed by developing simple mockup designs of the screen layouts. The output of this step helps in working with graphics, selecting appropriate icons for each screen, and meeting the required characteristics for user interaction. Many HCI aspects including metaphors and stakeholders [21] are incorporated in the app development in order to enhance its appeal from a user perspective.

C. Deployment of Relevant Software

The software tools we deploy in the development of the visualization tool (that also help to provide information for app development) are briefly explained herewith.

ArcGIS: The tool called ArcGIS serves as a geographic information system (GIS) software [22]. It allows handling and interpreting of geographic information by visualizing geographical statistics through layer building maps such as climate data or trade flows. It helps in data geo-processing as elaborated in the next subsection.

Tableau: Tableau [23] is a data visualization tool focused on aiding people to observe and interpret data at-a-glance with interactive visual features. Tableau adheres to strict requirements in using data to solve problems. It converts raw data into a various visualization formats with robust graphics. Its features are as follows.

Worksheets with visualizations exploring various aspects of the data
Dashboard combining different worksheets embeddable into websites
Trend lines to learn correlations
Custom filters and zooming features
Identification of areas that need attention
Future predictions based on previous data

These features of Tableau motivate us to harness it within the development of our visualization tool.

D. Big Data Geo-processing

Hydro-climatic big data relevant to precipitation in Sub-Saharan Africa is provided to us by environmental scientists in our team. The original data of the order of TB is in NetCDF (Network Common Data Format) file format [24]. NetCDF files contain metrics on predictions for climatic parameters, and are subjected to geo-processing so that they can be used to design the visualization tool and the app (see Fig. 1).

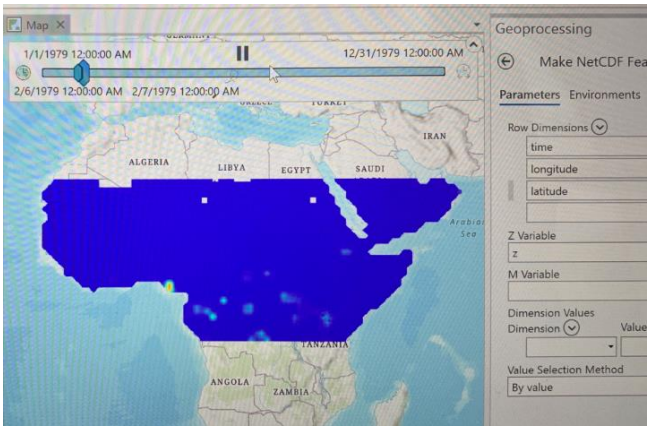


Fig. 1. ArcGIS used for Geo-processing of Sub-Saharan Preciptation Data

NetCDF is a file format used for storing multidimensional scientific data (variables) such as temperature, humidity, pressure, wind speed, and direction. Each of these variables can be displayed through a dimension (such as time), which is achieved using ArcGIS [22] by making a layer or table view from the corresponding NetCDF file. In order deploy Tableau, we need to transform the data from NetCDF to CSV (Comma Separated Values) format. ArcGIS is useful for this data conversion by providing geo-processing and mapping of hydro-climatic data (as found in some climate studies [25]).

The converted data has precipitation values per day for an individual geo-location identified by a unique latitude and longitude pair, stored in CSV files. These still contain big data in TB, further subjected to data reduction using classical techniques: numerosity reduction with stratified sampling, and dimensionality reduction with heuristic feature selection by a combination of forward feature selection and backward feature elimination. The resulting CSV data files are well-suited for analysis and provide the basis for data input in the visualization tool and app development.

IV. Precipation Data Visualization Tool

A data visualization tool for precipitation studies in Sub-Saharan Africa has been implemented using the pertinent big data containing metrics on climatic parameters, subjected to relevant preprocessing using ArcGIS for transformation, and classical techniques for numerosity as well as dimensionality reduction. In order to provide a customizable user experience, we embed the Tableau functionalities within our precipitation visualization tool. The tool is described here with respect to creating visualizations and analyzing interactive displays.

A. *Creating the Visualizations*

The steps involved in creating interactive data visualizations in our tool are as follows (see Fig. 2 on the tool home screen).

- Open the home screen in the tool, click on *File* from the menu bar and select *New* from the menu.
- From the left panel, select *Connect to Data* to link to the data source (in this case, the CSV files after data transformation and reduction)
- Once connected to the data, drag and drop the attributes in designated rows and columns.
- The right panel provides the user with a list of charts to select visualizations suited to the data.

Based on visualizations created, we delve into the interactive displays and their analysis.

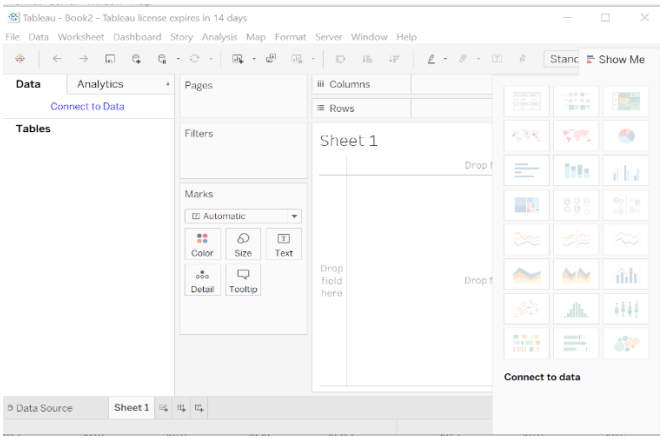


Fig. 2. Home Screen of the Tool for Preciptation Data Visualization

B. *Interactive Displays for Data Analysis*

We explain the interactive displays of our precipitation data visualization tool here, with dashboard screenshots of the respective features. These include precipitation trendlines, various filters, Sub-Saharan maps, zooming options etc.

a) Precipitation Trendlines: A trendline is an easily recognizable line drawn on a chart to connect a series of precipitation points together. The resulting line is then used to give an idea of the variation in precipitation. An example of a trendline is shown in Fig. 3 for Sub-Saharan data. The figure also illustrates its date filter option (top right). The slidebar allows users to select the date range and the trendline is displayed according to the selection. This feature enables at-a-glance analysis of preciptation in a given timeframe, thus allowing users to make decisions based on that, and envisage trends for similar future timeframes (months w.r.t years).

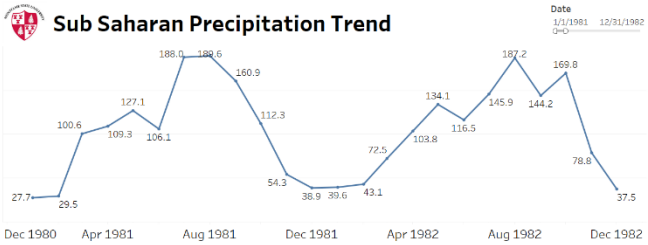


Fig. 3. Precipitation Trendline with Date Filter Option

b) Geographical Map Customization: The precipitation data visualization tool allows users to create and customize geographical maps with specific date ranges as well as latitude-longitude values. There are useful filters for this purpose such as date filter, and latitude-longitude filter.

c) Date Filter: The map customization feature of our tool provides a date filter that allows users to select dates based on a range, as shown in Fig. 4. Precipitation displayed on the map is altered based on the selected dates.

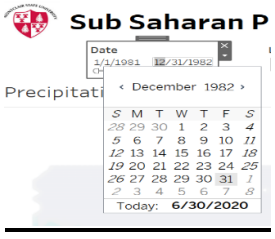


Fig. 4. Date Filter for Map Customization

d) Latitude-Longitude Filter: The tool provides a drop-down selection for latitude and longitude values in the customizable map. Each geographical location is uniquely identified by latitude and longitude coordinates, so this filter allows users to view precipitation accordingly (see Fig. 5).

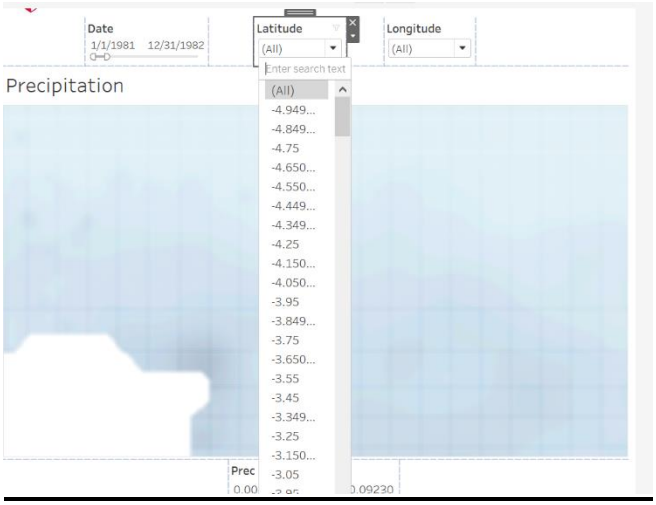


Fig. 5. Latitude and Longitude Filter for Maps

e) Zooming, Rectangle, Radial and Lasso Selection: There are various viewing options available in the tool to filter the dashboards and observe specific data points on the customizable maps. Zooming selection provides closer views of given regions. Rectangle selection provides views by marking and selecting within specific rectangular shapes. Radial selection offers views in given circular zones. Lasso selection allows users to draw freehand shapes around given regions and depicts the respective views. Fig. 6 illustrates *Lasso selection* (applied after *zooming* to get a closer view of a certain region). Such features allow users to connect regions to precipitation easily with options for customizing the views. These options thereby facilitate data analysis on precipitation based on regions and views.

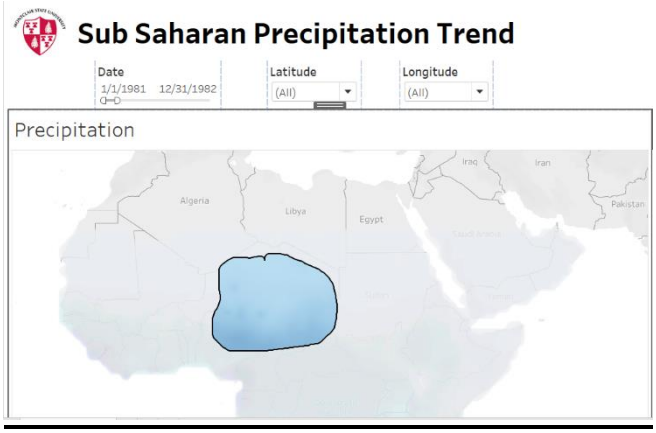


Fig. 6. Viewing Options (Lasso Selection)

f) Search By City: The precipitation data visualization tool offers a city filter through which users can type the name of any Sub-Saharan city, to view the precipitation trends for that particular city. This could include one specific city or all the cities in a given country. Likewise, various aspects of precipitation data can be studied using this tool.

V. Visualization Tool Evaluation and Discussion

The precipitation data visualization tool is subjected to user evaluation. Users are given the tool demo to evaluate it based on clarity, efficiency and usefulness. They are asked to assign scores on a Likert scale of 1 to 5 (1:lowest to 5:highest). In HCI, a Likert scale is a method often used for subjective to objective mapping to effectively assess user responses [21]. The averages of the evaluation scores appear in TABLE I. It is seen that all scores are high depicting user satisfaction.

TABLE I. Assessment: Preciptation Data visualization tool

User Evaluation of Visualization Tool	Likert Scale (1:lowest to 5:highest)		
	Clarity	*Efficiency*	*Usefulness*
Average Scores	4.50	4.37	4.75

The visualization for the climatic parameter of precipitation conducted in our tool intends to provide comprehensible information about inferences of rainfall. It specifically helps to fathom changes that can potentially occur in local regions. This tool is easily accessible and can be downloaded by public officials to advise the concerned users such as local farmers. It is also useful for environmental scientists to conduct studies on predictive analysis and sustainability.

The precipitation data visualization tool sets the stage for the design and implementation of a precipitation app for Sub-Saharan Africa. This is in order to further facilitate data access and analysis in a ubiquitous manner.

VI. Android App For Precipation Studies

A mobile application for the Android platform is developed for precipitation studies in Sub-Saharan Africa using the hydro-climatic big data provided by environmental scientists. The geo-processing conducted on the big data during the design of our visualization tool is useful, the transformed data being used in app development. The analysis conducted using this precipitation data visualization tool is also helpful in the design of the app, e.g. in creating the relevant screen layouts and icons. Principles from HCI [21] are incorporated in the app design in order to enrich the user experience.

A. App Mockup Design

In HCI, the concept of mockup design is crucial in order to grab user attention towards the app [26] since it offers a glimpse into the initial appearance of the complete app, and the users' requirements can be edited at this stage if needed. A few free and easily accessible web UI (user interface) Mockup tools are Balsamiq mockups, Mockplus, Mockplus iDoc, Wireframe.cc and Moqups [27].

Fig.7. App Layouts - Mockup Designs

Among these tools, the Balsamiq mockup tool [27] is used in our app to create the layouts of multiple screens (see Fig. 7). We prefer this tool due to its compatibility with our work. The design layouts are as follows.

*a) Layout1:*This UI provides the title of the app; its images reflect that it offers precipitation studies.

b) Layout 2: This UI is designed to provide the overall description of this app.

c) Layout 3: This UI serves the purpose of user input.

d) Layout 4: This UI displays the result in terms of minimum, maximum and average precipitation.

Note that in our app design, we incorporate HCI aspects such as stakeholders, metaphors, Fitt's Law and ethnographic study [21, 28]. *Stakeholders* in HCI refer to significant people involved with the development and the use of the app. In our context these include domain experts from environmental science, software engineers and computer scientists, as well as naïve users. Accordingly, our mockup designs are thus subjected to feedback from such stakeholders. A *metaphor* in HCI [28] is a pictorial depiction of a concept (e.g. a bird for the Twitter site). We use metaphors such as clouds and droplets to depict rainfall, and friendly icons for the opening welcome screen. These are appreciated by stakeholders as per their feedback. The classic *Fitt's Law* [28] states that the time to access an object is directly proportional to its size and inversely proportional to its distance from the start screen. Thus, we have designed objects (metaphors, buttons, entry fields etc.) to be large enough for users to spot easily, yet small enough to fit on a screen. We have kept navigation simple such that all objects are fairly near the opening screen, and thus, users can quickly find them. Another aspect of HCI is *ethnographic study* [28] which refers to detailed knowledge of the concerned domain based on observation and analysis. We conduct this study via comprehension of Sub-Saharan Africa with respect to its geography and specific needs. In addition, we incorporate the knowledge of domain experts to guide us based on ethnographic details. Since these aspects are incorporated in our mockup design, stakeholders really appreciate it, accepting the design as it is. Thus, we use this design in the actual layouts of the app for implementation.

B. Database Development for the App

Based on our understanding of the data from the precipitation visualization tool, the converted CSV data is still huge by app development standards (although acceptable for the tool). Hence, the SQLite database system [29] is used to further reduce the data representation by using non-parametric data cube aggregation techniques. Excel is used to further analyze the data and make it ready for use in the app's SQLite database. The data provides precipitation per day for an individual geo-location identified by a unique pair of latitude and longitude. Fig. 8 illustrates this process of Database (DB) development and for the precipitation app.

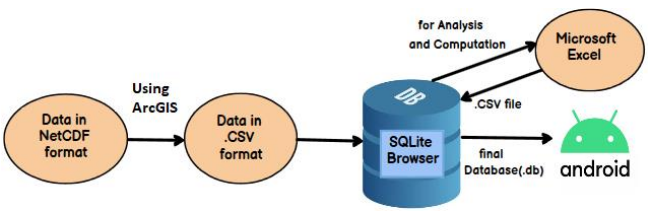


Fig. 8. DB Development for the Precipitation App

C. App Implementation Details

In current technological era, smartphones are an integral part of our lives. A recent US Mobile App Report from comScore indicates that 50% of us spend our digital media time in smartphone apps [30]. Due to the broad price range from very low to high end, Android users span the globe, including lower income areas and developing nations. Hence, we choose the Android platform for our app development, especially since we target Sub-Saharan Africa. The app implementation with the respective tools occurs as follows.

a) Android Studio: This is an official Integrated Development Environment (IDE) for Android apps [31]. In Android Studio, apps are considered to be a project with source code and resource files. There are multiple folders inside the project. The Manifest folder contains a file called AndroidManifest.xml, the Java folder contains multiple Java source code files and the Res folder contains XML layout code, UI strings and bitmap images (see Fig. 9). All the build files are visible at the top level under Gradle Scripts.

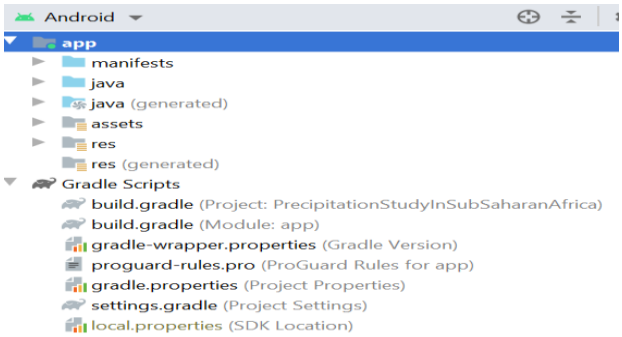


Fig. 9. App Structure in the Android Studio Environment

b) Android Emulator: The Emulator is an inbuilt feature of the Android Studio, used to test and run the app on different API level devices [32]. It provides almost all the capabilities of a real Android device. Testing the app frequently after writing or editing codes is relatively easier on the Emulator than on a real Android device. It comes with predefined configurations for various Android phones, tablet, wear OS and Android TV Devices.

c) SQLite Database: As stated earlier, we create the DB for this app using SQLite. This is an open source software library that provides a relational database management system. It stores data to a text file on the device. The Android Studio has an inbuilt SQLite implementation and is available in the "android.database.sqlite" package [29]. Based on this, Fig. 10 is a part of the Java code to open the DB in SQLite.

DatabaseOpenHelper.java

```
package com.sudha.precipitationstudyinsubsaharanafrica;

import android.content.Context;
import android.database.sqlite.SQLiteDatabase;
import com.readystatesoftware.sqliteasset.SQLiteAssetHelper;

public class DatabaseOpenHelper extends SQLiteAssetHelper {

    public static final String DATABASE_NAME="PrecipitationDatanew.db";

    //constructor
    public DatabaseOpenHelper(Context context) {
        super(context, DATABASE_NAME, factory: null, version: 1);
    }
}
```

Fig. 10. Relevant Snapshot of Java Code for Opening the DB in SQLite

d) Android Geocoder: Android has an inbuilt geocoder API available in "android.location.Geocoder". Geocoding is

the process of converting the address to its corresponding latitude and longitude value. In order to use geocoder, the value of the actual API key in the manifest file is required.

```
try {
    TextView addressTextView = findViewById(R.id.editTextAddress);
    String address = addressTextView.getText().toString();
    Geocoder geocoder = new Geocoder(context);
    List<Address> addresses = geocoder.getFromLocationName(address,
1);
    if(!addresses.isEmpty()) {
        Address geoAddress = addresses.get(0);
        String latitude = Double.toString(geoAddress.getLatitude());
        String longitude = Double.toString(0 +
geoAddress.getLongitude());
```

Fig. 11. Java Code Snippet for Geocoder Conversion

The relevant Java code snippet that performs the geocoder conversion in our app is shown in Fig. 11 here.

D. App Execution and Data Flow

Considering the implementation described herewith, the steps in the execution of our Android app along with data flow based on inputs and outputs are listed next.

- The Android inbuilt functionalities such as code editing, debugging and profiling are used.
- The already created SQLite DB is imported into the asset folder inside the Java folder.
- App UI (user interface) is created using relative and linear layouts. Its outlook is enhanced by adding background images, user-friendly icons and buttons.
- Geocoding API is used to get latitude and longitude of the locations entered in the user input.
- The converted latitude and longitude with dates are used to query for precipitation from the DB.
- If user-entered values match with the DB, then the corresponding precipitation values along with min, max, average and total precipitations are displayed on app else "No data found" is displayed.

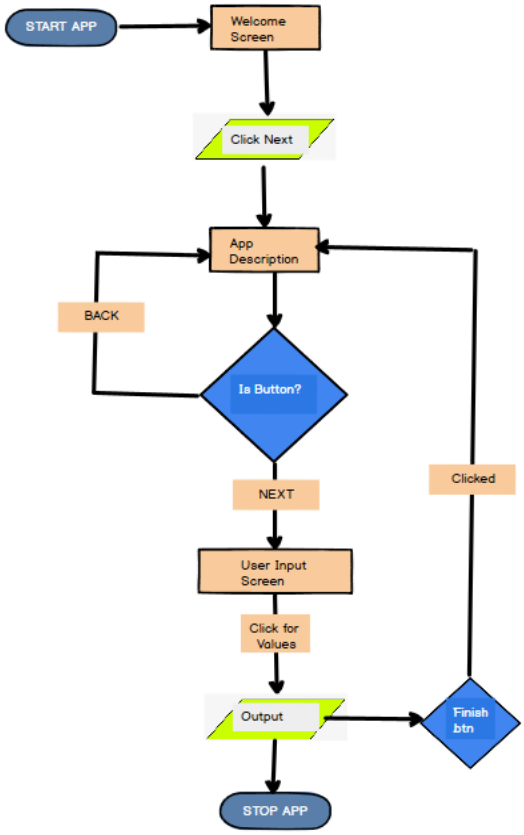


Fig. 12. App Execution Flowchart Diagram

Based on this execution, the flowchart of the app is shown in Fig. 12. This depicts the working and data flow of the app.

VII. APP EVALUATION AND DISCUSSION

The app has undergone evaluation by targeted users to assess it for precipitation studies in Sub-Saharan Africa. Fig. 13 is a snapshot of app UI layouts as per the app design presented to users. Fig. 14 shows an example of inputs entered by the users and corresponding outputs depicted by the app.

Based on this, objective scores are obtained by asking users to evaluate the app for its clarity, efficiency and usefulness (analogous to the visualization tool). The same HCI-based Likert scale of 1 to 5 (1:lowest to 5:highest) is used here. The results of user evaluation are summarized in TABLE II. High scores in all categories (based on average of scores assigned by users) indicate the overall positive reception of the app.

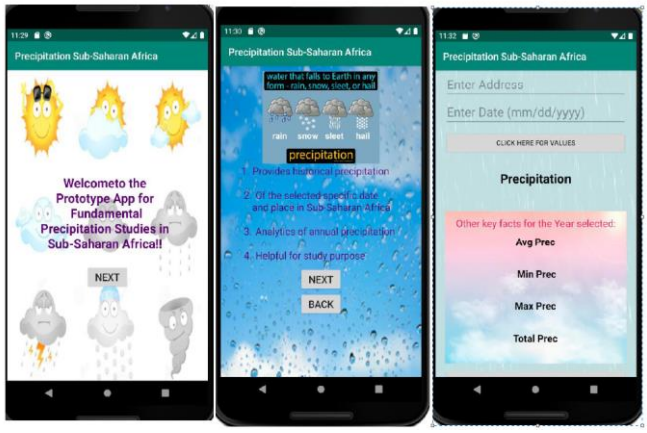


Fig. 13. App UI Layouts Subjected to User Evaluation

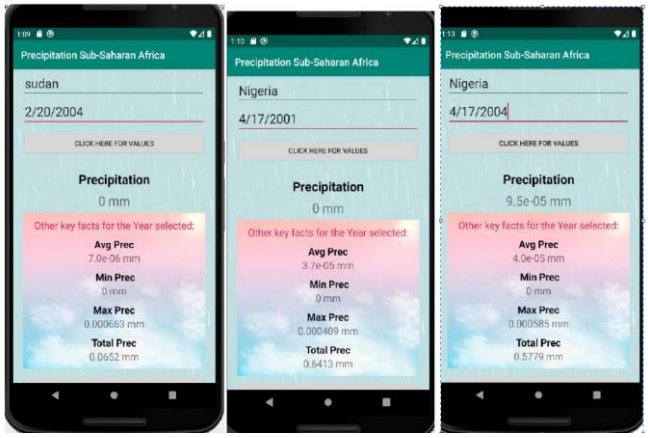


Fig. 14. User Inputs and Corresponding App Outputs

TABLE II. ASSESSMENT: PRECIPTATION MOBILE APPLICATION (APP)

User Evaluation of Mobile Application	Likert Scale (1:lowest to 5:highest)		
	Clarity	*Efficiency*	*Usefulness*
Average Scores	4.22	4.68	4.90

Based on the evaluation, we can infer that this app is suitable for use and can be helpful to wide range of users. Climate change is a huge environmental concern today as evident from a variety of studies, surveyed recently [33]. Over 2 billion people derive their livelihood from agriculture. Rain-fed agriculture is a type of farming that relies on rainfall for

water and accounts for more than 95% of farmed land in Sub-Saharan Africa, 90% in Latin America, 75% in the Near East and North Africa, 65% in East Asia and 60% in South Asia [34]. Studies on precipitation can be indispensable sources of support to the rain-fed agriculture so that adequate planning can be done for crop cultivation. In line with such work, our easily accessible and downloadable Sub-Saharan African precipitation app can offer localized information to farmers and other residents to plan daily activities and make key decisions. We have linked relevant app files to GitHub and these can be available to some interested users upon request.

VIII. Conclusions and Future Work

In this paper, we address precipitation studies in Sub-Saharan Africa. This work makes the following contributions:

- Developing a tool for at-a-glance visualization and analysis for precipitation data
- Building an app that provides ubiquitous access to precipitation data for users on mobile devices
- Conducting evaluation of the tool and app with targeted users and obtaining positive reception.
- Providing useful inputs to environmental scientists, public agencies, farmers and local residents
- Laying the concept of a visualization tool and app that can be extended to other environmental studies

Our work in this paper can set the stage for predictive analysis using Artificial Intelligence, especially Machine Learning techniques for studies on hydro-climatic data. While we have only considered precipitation as the parameter in this work, which might be construed as a limitation, these studies can be seamlessly extended to cover other climatic parameters such as temperature and humidity. Also, the methods and software described here for Sub-Saharan Africa can be used for other geographic areas with suitable adaptation. This helps in sustainable living on the whole. Our work thus makes broader impacts on sustainability in addition to direct impacts on the realms of HCI, IoT, Big Data and Visualization.

References

[1] T. Didiano, M. Johnson, T. Duval, "Disentangling the Effects of Precipitation Amount and Frequency on the Performance of 14 Grassland Species", PLOS ONE, 2016, 11(9): e0162310.

[2] M. Kirschbaum, A. Mcmillan, "Warming and Elevated CO2 Have Opposing Influences on Transpiration. Which is more Important?". Current Forestry Reports Journal, 2018, 1(21):10.

[3] S. Sellars, P. Nguyen, W. Chu, X. Gao, K. Hsu, S. Sorooshian, "Computational Earth Science: Big Data Transformed Into Insight", EOS Transactions. AGU, 2013, 94(32):277.

[4] Y. Chaowei, H. Qunying, L. Zhenlong, L. Kai, H. Fei, "Big Data and Cloud Computing: Innovation Opportunities and Challenges", International Journal of Digital Earth, 2017, 10(1):13-53.

[5] S. Contractor, M. Donat, L. Alexander, M. Ziese, A. Meyer-Christoffer, U. Schneider, E. Rustemeier, A. Becker, I. Durre, R. Vose R. "Rainfall Estimates on a Gridded Network (REGEN) – A Global Land-Based Gridded Dataset of Daily Precipitation from 1950–2013", Hydrology and Earth System Sciences Discussions, 2019, pp. 1-30.

[6] A. Hoscilo, H. Balzter, E. Bartholomé, M. Boschetti, P. Brivio, A. Brink, M. Clerici, J. Pekel, "A Conceptual Model for Assessing Rainfall and Vegetation Trends in sub-Saharan Africa from Satellite Data, International Journal of Climatology", 2014, 35:10.1002.

[7] S. Barrios, E. Strobl, L. Bertinelli, "Climatic Change and Rural–Urban Migration: The Case of Sub-Saharan Africa", Journal of Urban Economics, 2006, 60: 357-371.

[8] J. Juana, Z. Kahaka, F. Orukut, "Farmers' Perceptions and Adaptations to Climate Change in Sub-Sahara Africa: A Synthesis of Empirical Studies and Implications for Public Policy in African Agriculture". Journal of Agricultural Science, 2013, 5:121-135.

[9] M. Muller, "Adapting to Climate Change: Water Management for Urban Resilience, Environment & Urbanization", 2007, 19(1): 99–113.

[10] D. Pathak, A. Varde, C. Alo, F. Oteng, "Ubiquitous Access for Local Water Management through HCI-Based App Development", IEEE UEMCON 2019, pp. 226-232.

[11] C. Varghese, A. Varde, X. Du, "An Ordinance Tweet Mining App to Disseminate Urban Policy Knowledge for Smart Governance", I3E conference 2020, vol. 2, pp. 389-401.

[12] M. Puri, A. Varde, X. Du, G. de Melo, "Smart Governance Through Opinion Mining of Public Reactions on Ordinances", IEEE ICTAI, 2018, pp. 838-845.

[13] X. Du, A. Varde, R. Taylor, "Mining Ordinance Data from the Web for Smart City Development", DMIN conference, 2016, ISBN: 1-60132-453-7, CSREA Press.

[14] X. Du, A. Varde, "Mining Multicity Urban Data for Sustainable Population Relocation", Intl. J. of Computer, Electrical, Automation, Control and Information Engineering, 2016, 9(12):2441-2448.

[15] A. Pampoore-Thampi, A. Varde, D. Yu, "Mining GIS Data to Predict Urban Sprawl", ACM KDD Bloomberg Track, 2014, pp. 118-125.

[16] M. Pawlish, A. Varde, S. Robila, "Decision Support in Data Centers for Sustainability", IEEE ICDM workshops 2013, pp.613-620.

[17] M. Pawlish, A. Varde, A. Robila, C. Alvarez, C. Fleischl, G. Serviano, "GreenDSS Tool for Data Center Management", IEEE Intl. Conf. on Information and Communication Systems, 2014, pp. 1-6.

[18] D. Bacic, A. Fadlalla, "Business Information Visualization Intellectual Contributions: An Integrative Framework of Visualization Capabilities and Dimensions of Visual Intelligence", Decision Support Systems, 2016, 89:10.1016/j.dss.2016.06.011.

[19] L. Perkhofer, C. Walchshofer, P. Hofer, "Does Design Matter when Visualizing Big Data? An Empirical Study to Investigate the Effect of Visualization Type and Interaction Use", Journal of Management Control, 2020, 31:55-95.

[20] R. Benestad, "Downscaling Climate Information", Oxford Research Encyclopedia of Climate Science, Oxford University Press, 2016.

[21] S. Thuseethan, S. Kuhanesan, "Effective Use of Human Computer Interaction in Digital Academic Supportive Devices", International Journal of Science and Research (IJSR), 2014, pp. 388-392.

[22] H. Singh, S. Bawa S, "Spatial Data Analysis with ArcGIS and MapReduce", Intl. Conf. on Computing, Communication and Automation, 2016, pp. 45-49.

[23] Tableau Software, https://www.tableau.com/products/desktop

[24] R. Rew, G. Davis, "NetCDF: An Interface for Scientific Data Access, IEEE Computer Graphics and Applications", 1990, 10(4):76-82.

[25] L. Cattaneo, V. Rillo, P. Mercogliano, "Clime: Analysing Climate Data in GIS Environment", SISC conference, 2014, pp. 434-447.

[26] Mockup Design, https://www.buildinary.com/news/the-benefit-of-app-design-mockups

[27] Mockup Tools, https://medium.com/@tristaljing/4-best-web-ui-mockup-tools-for-free-89a1513c3fcd

[28] Y. Rogers, H. Sharp, J. Preece, Interaction Design: Beyond Human-Computer Interaction, 2015, 4th Edition,Wiley, 978-1-119-02075-2.

[29] SQLite Database, https://www.sqlite.org/index.html

[30] Y. Liu, F. Li, L. Guo, B. Shen, S. Chen, "A Comparative Study of Android and iOS for Accessing Internet Streaming Services", Intl. Conf. on Passive and Active Measurement, 2013, pp. 104-114.

[31] Android Studio, https://developer.android.com/studio/intro

[32] Android Emulator, https://developer.android.com/studio/run/emulator

[33] X. Du, M. Kowalski, A. Varde, G. de Melo, R. Taylor, "Public Opinion Matters: Mining Social Media Text for Environmental Management", ACM SIGWEB, 2019, 5(1):5-15.

[34] O. Serdeczny, S. Adams, F. Baarsch, D. Cuomou, A. Robinson, B. Hare, M. Schaeffer, M. Perrette, J. Reinhardt, "Climate Change Impacts in Sub-Saharan Africa: From Physical Changes to their Social Repercussions", Regional Env. Change, 2016, ISSN: 1436 – 3798.

Diagnostic Classification Using Hepatitis C Tests

Sudarshan S. Chawathe
School of Computing and Information Science & Climate Change Institute
University of Maine
Orono, Maine 04469-5711, USA
chaw@eip10.org

***Abstract*—This paper describes methods for automated classification of individuals by Hepatitis C medical category using data from a series of commonly used diagnostic tests. The methods are evaluated experimentally using a publicly available dataset from prior work. The accuracy of some methods compares favorably with similar results reported in prior work. In addition to quantitative results on prediction accuracy, training and testing times, and model sizes, the paper includes a detailed look at some concrete representative classifiers generated by a few of the competitive methods, permitting a human domain expert to further study the models and classifiers.**

***Index Terms*—Medical Informatics; Classification.**

I. Introduction

Hepatitis C is a significant worldwide health concern, with the World Health Organization estimating 71 million cases in 2017. It is a viral disease for which there is no vaccine as yet, making prevention of its spread even more crucial. Such prevention is significantly aided by detection based on blood tests and other variables.

The *main question addressed by this paper* is: How well can automated methods of analyzing laboratory tests and related variables for Hepatitis C perform, both quantitatively (accuracy, running times, etc.) and qualitatively (predictions that amenable to analysis by a human expert)? The paper adopts an experimental approach based on a publicly available dataset from prior work [1] and studies the problem not only quantitatively but also qualitatively in order to gain insight into the characteristics of the underlying classification problem and biomedical variables.

The main *contributions* may be summarized as follows: This paper

- augments a publicly available dataset with descriptions that make it more accessible to further use by others.
- studies the effectiveness of a diverse panel of well established classification algorithms in determining the Hepatitis C category from a panel of commonly used lab test results and related variables.
- quantifies the classification merit of each of the attributes to study whether comparable results are achievable with fewer attributes.
- summarizes experimental results quantifying the accuracy and performance of classification methods using both the full set of attributes and a reduced one with only three attributes.
- includes detailed descriptions of the classifiers (rules, trees, etc.) in order to permit validation of some of the results by a human expert.

The main *results* may be summarized as follows:

- High classification accuracy is achievable using a few different classification algorithms. In addition to the often well performing ensemble-based Random Forest classifier, these methods include non-ensemble methods such as a simple unordered rules classifier that uses rough sets (fuzzy rules).
- An examination of the details of the rule and tree based classifiers permits a human expert to validate them and to use them to guide further research.
- The Bayes Net classifier not only provides high accuracy but also provides a nontrivial network that would benefit from further study.
- The quantitative results are limited by the modest amount of data and further by the very skewed distribution across categories and further work with larger datasets will be valuable.

Paper outline: The classification problem in the context of Hepatitis C categories and the attributes from an array of tests is presented in Section II based on the dataset due to prior work [1]. The primary experimental setup and results are described in Section III using both the full dataset and a version with a reduced number of attributes. The selection of the attributes in the latter set is addressed in the more general context of attribute merit in Section IV. A representative selection of concrete classifiers generated by the methods found to be competitive in Section III is examined in more detail in Section V. Some related work is described in Section VI, and Section VII summarizes the paper and outlines topics for future work.

II. Data from Tests

The dataset produced by prior work [1] has instances classified into one of five Hepatitis C categories as summarized by Table I. The other attributes in this dataset are identified by the names listed in the second column of Table II. That table also includes the abbreviations used in the rest of this paper as well as the brief descriptions for reference, from other sources. Apart from the category and the two integer-coded attributes (age and sex), all other attributes are real numbers.

The main task is that of predicting the category using the other attributes in Table II. However, as described later, this

978-1-7281-9615-2/20/$31.00

TABLE I
HEPATITIS C CATEGORIES IN ORIGINAL DATASET [1] AND THEIR CODES.

code	Hepatitis C category
0	blood donor
4	suspect blood donor (*0s* in original dataset)
1	Hepatitis
2	Fibrosis
3	Cirrhosis

TABLE II
ATTRIBUTE NAMES FROM THE HEPATITIS C DATASET [1], WITH ABBREVIATIONS AND DESCRIPTIONS ADDED.

abbr.	name	description
age	Age	participant's age in years (integer)
sex	Sex	participant's sex (0/1/2 for female/male/other)
alb	ALB	albumin
alp	ALP	alkaline phosphatase
alt	ALT	alanine amino-transferase
ast	AST	aspartate amino-transferase
bil	BIL	bilirubin
che	CHE	choline esterase
cho	CHOL	choline
cre	CREA	creatinine
ggt	GGT	γ-glutamyl-transferase
prt	PROT	total protein
cat	Category	Hepatitis C category

paper also studies the task of performing that prediction using a much smaller set of attributes. The distribution of the data across categories and sex is worth noting and is depicted by Fig. 1 because, in particular, it is very skewed. (Note the logarithmic scale on the vertical axis of that chart.)

III. EXPERIMENTAL EVALUATION

For the purposes of experimental evaluation, a system for data preparation and classification was implemented using the Kawa Scheme programming environment which runs on the Java Virtual Machine (JVM) platform. SQLite [2] is used as a relational data store for the data imported from its raw CSV (comma-separated values) form after some normalization. The Java API to the WEKA workbench [3] is used from Kawa Scheme to access the former's implementations of various classification algorithms. The experiments were conducted on a notebook computer with an Intel Core i5-5250U CPU (2 cores, 4 threads, 3 MB cache, 5 GT/s bus) clocked at 1.6 GHz (2.7 GHz max. turbo), 8 GiB of DDR3 RAM, and an SSD for secondary storage. All the experiments use conventional 10-fold cross validation for evaluation and include multiple trials. The error-bars in the charts summarizing the results represent the standard error of the mean.

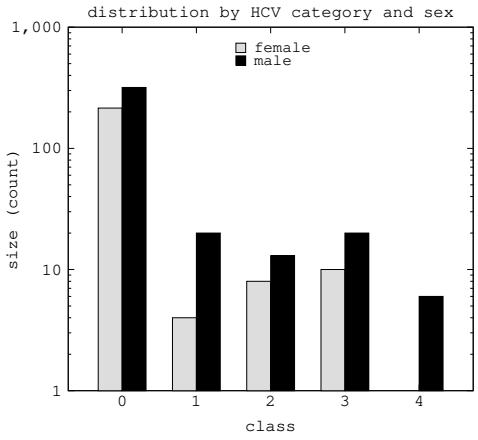


Fig. 1. Distribution of instances across HCV categories by sex, using a logarithmic scale on the vertical axis.

The choice of classification algorithms was guided by the desire to include diverse methods (e.g., ensemble, lazy, rules) as well as by the results of some preliminary experiments to identify the promising candidates. The experimental set includes 0R (`ZeroR`) as a baseline method, which is particularly important for this dataset because of the large skew in class distributions. (The parenthesized abbreviations are used in the presentation of the experimental results.) It includes several rules-based classifiers due to their benefit of human understandability: 1R (`OneR`) to serve as a baseline within this group, Decision Tables (`DTab`), PART rules (`PART`), RIPPER-style rules [4] (`JRip`), and the FURIA [5] fuzzy unordered-rules classifier (`FURI`). Tree-based classifiers are represented by C4.5-style trees (`J48`), Random Tree (`RTre`), and Random Forest (`RFor`), the last of which is also an ensemble method. Function-based classifiers are represented by the simple logistic (`SLog`) and logistic (`Logi`) methods as well as sequential minimal optimization (`SMOp`) for support vector machines (SVM). Bayesian classifiers are represented by Naive Bayes (`NBay`) and two versions of Bayes Net: one with a single parent per node (`BNt1`) and the other an unlimited number of parents per node (`BNtu`). Finally, lazy or instance-based methods are represented by k-nearest neighbor (`IBkN`), locally weighted learning (`LWtL`), and K-star (`KStr`).

The methods are evaluated using two datasets. The first, denoted `full`, includes all the attributes from Table II. The original dataset also includes a unique identifier per instance. However, preliminary experiments revealed a strong correlation between that identifier and the category labels. Since such a correlation is unrealistic in practice, the identifier was excluded from the datasets used in these experiments. The second dataset, denoted `top3`, includes only the most promising attributes as quantified by the methods of Section IV: `alt`, `ast`, and `che`. The purpose of using this dataset is to gain an understanding of the trade-off between prediction accuracy and cost (both computational, as quantified here, and the more significant cost of the underlying medical tests).

Fig. 2 summarizes the accuracy of each of these methods on both the `full` and `top3` datasets. It is notable that the Bayes Net methods provide accuracy superior to that of the usually proficient Random Forest method. As well, the much simpler rule-based method FURIA also provides competitive accuracy. Therefore some qualitative details of the Bayes Net and FURIA methods are included below, along

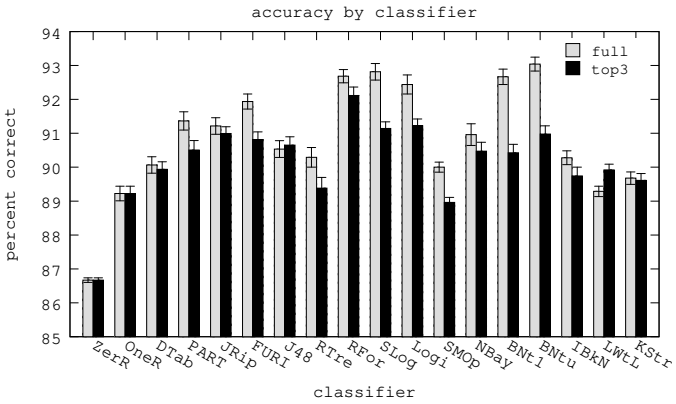


Fig. 2. Percentage of instances correctly classified by method, using all attributes and only three attributes. Here and in the charts that follow, the error bars mark standard error of the mean. In this chart and in others that follow, the error-bars, sometimes too small to discern, represent the standard error of the mean.

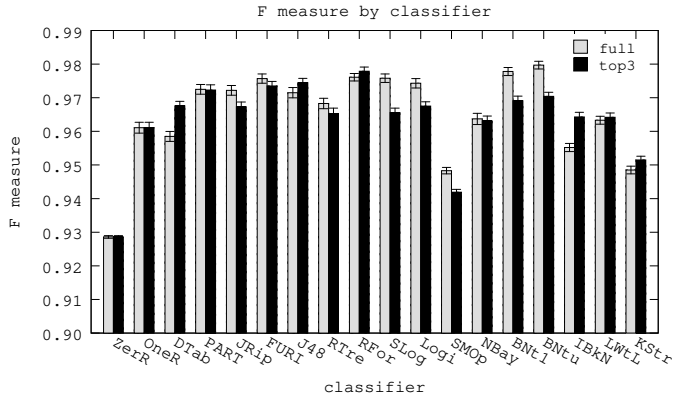


Fig. 3. Classification F-measure by method.

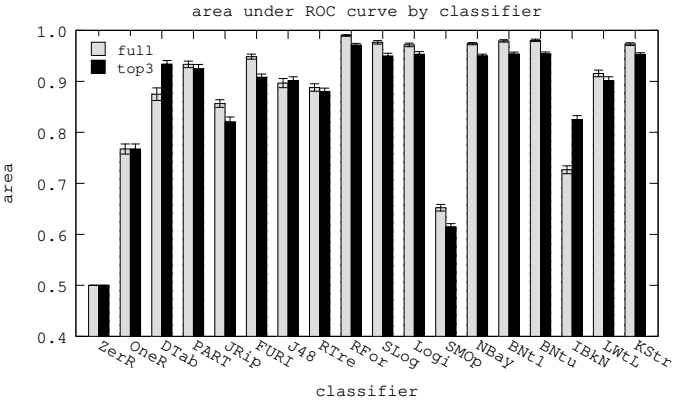


Fig. 4. Area under the ROC (receiver operating characteristic) curve by method.

with similar results for J48 for comparison. Given the highly skewed distribution of the data as summarized by Fig. 1, it is important to use a metrics such the F measure and area under the ROC (receiver operating characteristic) curve when quantifying accuracy and the results are summarized by Fig. 3 and Fig. 4, respectively. These results provide further evidence for the strengths of the Bayesian methods as well as FURIA for this problem.

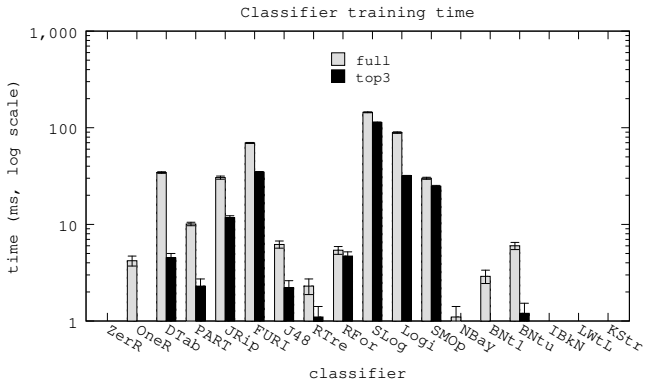


Fig. 5. Training (model building) time by method, using a logarithmic scale on the vertical axis.

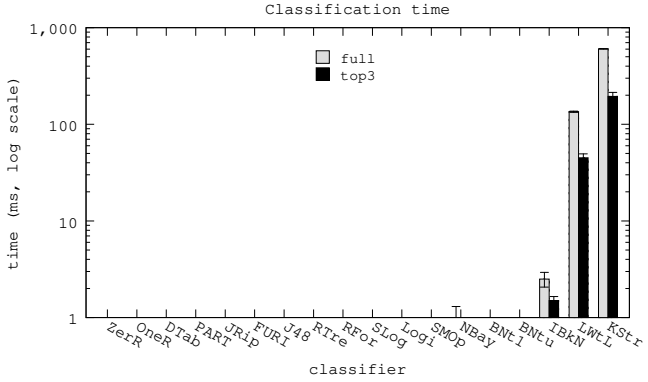


Fig. 6. Testing (classification) time by method, using a logarithmic scale on the vertical axis.

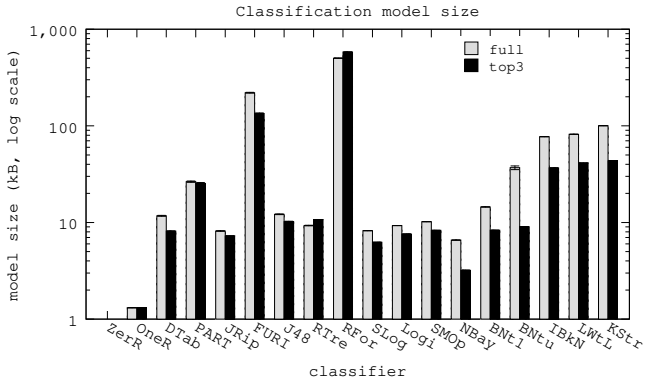


Fig. 7. Learned model size by method, using a logarithmic scale on the vertical axis.

Fig. 5 summarizes the time required to train the classification models. There is a considerable variation in these times across the methods, as highlighted by the logarithmic scale used on the vertical axis. The near-zero times for the three lazy methods listed last are not surprising. However, there are also considerable differences among the other methods. These numerical results on running times require some caution in interpretation in that they are dependent on the specific implementations of the underlying algorithms in WEKA as used by the prototype implementation. Fig. 6 summarizes similar results for the classification or testing times. The main observation here is that, except for the lazy methods, all testing times are too low (sub-millisecond) to meaningfully measure and low enough to not be of practical concern. There is substantial variation among the running times for the lazy methods but here too the absolute values are small enough to not be of much practical concern. Finally, Fig. 7 summarizes the serialized sizes of the models learned by different methods on both datasets.

IV. Attribute Merit and Selection

It is worthwhile to study the merit of different attributes for the purposes of predicting the Hepatitis C Category because such study provides a better understanding of the trade-off between the potentially higher accuracy resulting from a larger set of attributes and the lower costs of smaller sets of attributes. These costs include not only the computational costs as briefly quantified in the previous section, but also the much more significant costs (both financial as well as potential patient discomfort) of the underlying medical tests.

The first set of experiments in this section is based on intrinsic metrics for attribute merit (independent of specific classification methods): The results for the *Information Gain*, *Symmetric Uncertainty*, and *Gain Ratio* metrics are summarized by Fig. 8, and the results using Pearson's correlation are summarized by Fig. 9. There are some significant variations across metrics, but also some common features, such as the high merit of `ast`.

The second set of experiments in this section is based on the use of classifiers to gauge merit. As a representative selection, Fig. 10 summarizes the results for the 1R as a baseline and for FURIA and Bayes Net (with unlimited parents per node) as examples of classifiers with favorable accuracy and explainability characteristics (from Section III). Here as well we may observe the consistently high merit for some attributes, such as `ast` and `alt` (Table II).. The above observations are strengthened by experiments using correlation-based feature subset selection [6] to select a favorable set of attributes, which results in the set used for the `top3` dataset in Section III: `alt`, `alt`, and `che`.

V. Examining Classifiers

This section examines some specific classifiers generated by a few of the methods studied earlier for the purposes of determining the ease with which the classifiers' predictions may be interpreted by a human expert.

TABLE III
Confusion matrix for the 12-attribute Bayes Net classifier of Fig. 14.

category	predicted category				
	c0	c4	c1	c2	c3
c0	530	1	1	1	0
c4	0	6	0	1	0
c1	9	0	9	5	1
c2	5	0	7	9	0
c3	2	1	2	4	21

TABLE IV
Confusion matrix for the three-attribute Bayes Net classifier of Fig. 15.

category	predicted category				
	c0	c4	c1	c2	c3
c0	524	0	7	2	0
c4	4	0	1	2	0
c1	13	0	10	1	0
c2	5	1	8	5	2
c3	2	0	5	3	20

Based on the results of Section III, it is natural to begin by examining the classifiers generated by Bayes Net algorithm with unlimited parents per node (`BNtu`). The networks generated for the complete (`full`) and reduced (`top3`) datasets of Section III are depicted by Fig. 14 and Fig. 15, respectively, and the corresponding confusion matrices appear in Table III and Table IV. It is notable that there is nontrivial structure in the learned network. Determining whether this structure is significant from the biomedical perspective is a task for experts in that field, potentially guided by such results.

Although the fuzzy (rough sets) unordered-rules classifiers produced by the FURIA method have accuracy significantly lower than those from Bayes Net (Figs. 2 to 4), they are worth studying due to their simplicity and ease of interpretation. A sample classifier from FURIA for the full dataset appears in Fig. 13. The classifier for the `top3` dataset is very similar and is omitted for brevity.

TABLE V
Confusion matrix for the tree classifier of Fig. 11.

category	predicted category				
	c0	c4	c1	c2	c3
c0	524	1	2	6	0
c4	3	0	0	1	3
c1	10	0	8	3	3
c2	7	0	5	7	2
c3	2	1	1	5	21

Finally, even though the RIPPER-style J48 classifiers are not as competitive as the above for this problem, they are worth studying because in addition to of their simplicity and ease of understanding, they are also very well understood and widely implemented. The classification trees generated for the full and reduced datasets appear in Fig. 11 and Fig. 12, respectively,

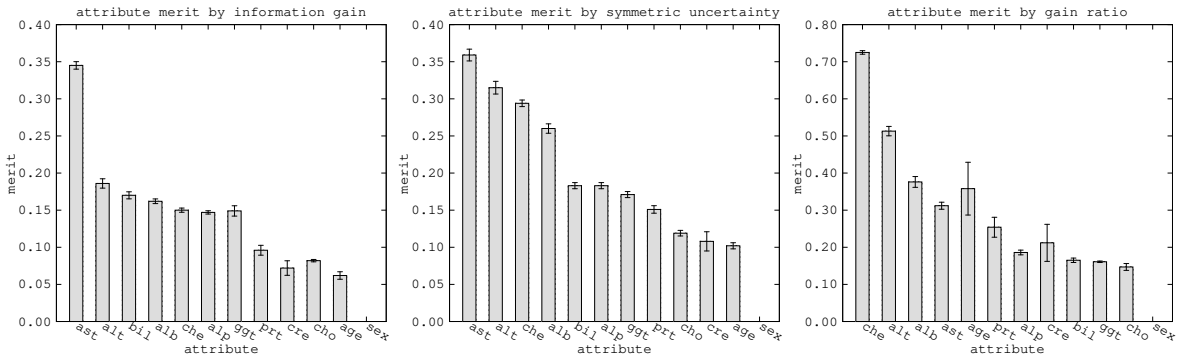


Fig. 8. Attributes (Table II) ranked by their merit by the *Information Gain*, *Symmetric Uncertainty*, and *Gain Ratio* metrics.

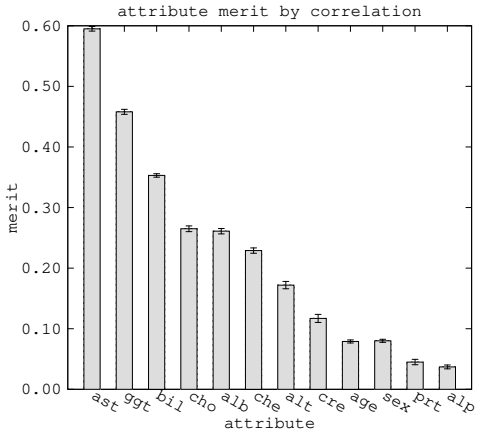


Fig. 9. Attribute merit by Pearson's correlation. Attribute descriptions appear in Table II.

TABLE VI
CONFUSION MATRIX FOR THE THREE-ATTRIBUTE TREE CLASSIFIER OF FIG. 12.

category	predicted category				
	c0	c4	c1	c2	c3
c0	524	0	5	1	3
c4	4	0	1	1	1
c1	10	1	3	6	4
c2	5	0	5	8	3
c3	1	0	6	3	20

TABLE VII
CONFUSION MATRIX FOR THE UNORDERED RULES, THREE-ATTRIBUTE CLASSIFIER OF FIG. 13.

category	predicted category				
	c0	c4	c1	c2	c3
c0	526	0	1	3	3
c4	4	0	0	2	1
c1	10	0	4	7	3
c2	4	1	5	7	4
c3	3	0	1	3	23

and the corresponding confusion matrices appear in Table VI and Table VI.

VI. RELATED WORK

The algorithms underlying the classification methods used in this paper are very general and have found very diverse applications, such as solving linear systems and performance modeling [8]. Although the dataset used in this paper focuses on blood tests, the classification study could be profitably extended to a more diverse array of tests enabled by the rapidly evolving sensor technologies [9]. In some such applications, rapid and reliable transmission of test data as enabled by some recent work [10]–[12] is a significant consideration as well. Determining and modeling contact among individuals based on diverse video processing techniques [13] is a complementary aspect of predicting and controlling the transmission of infectious diseases, especially those that are easily transmitted by brief contact. In the broader context of public health records and informatics, there is an ever growing need for robust data centers and related infrastructure as studied by some recent work [14].

The implementation used by the experiments benefits from the rapid prototyping on the JVM platform as enabled by the Kawa Scheme system [15], together with database work in SQLite [2], and from the well exercised implementations of classification methods provided by the WEKA workbench [3].

VII. CONCLUSION

This paper has examined automated classification of individuals into Hepatitis C categories based on data from a widely used panel of tests. Based on experiments with a publicly available dataset from prior work, has summarized quantitative results on accuracy and performance. As well, it has provided a detailed look at some concrete classifiers generated by the promising methods in order to permit insights into their suitability based on input from human domain experts.

The results are promising quantitatively although limited by the characteristics of the dataset (size and skew) and extending them to larger and more diverse datasets is a topic for ongoing and future work. Likewise, the qualitative results, such as the networks of Fig. 14 and Fig. 15, are interesting and would benefit from attention by domain experts.

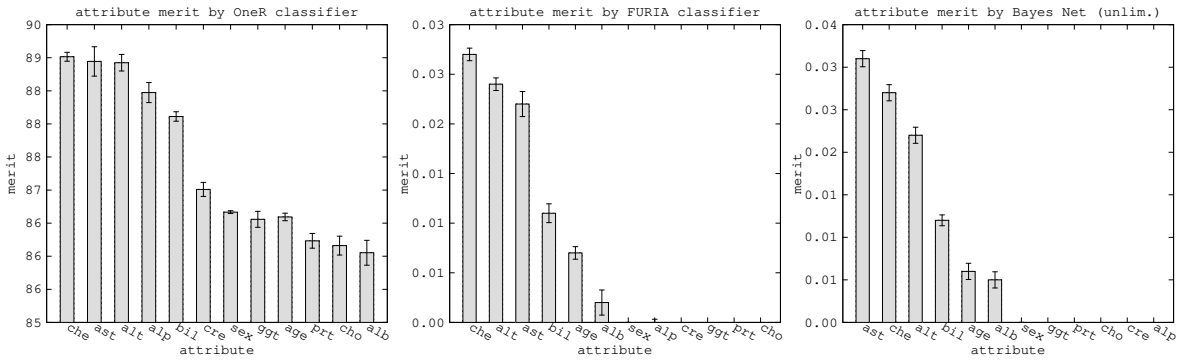


Fig. 10. Attribute merit by the classifiers: OneR (left), FURIA (middle), and Bayes Net with unlimited parents per node (right). Attribute descriptions appear in Table II.

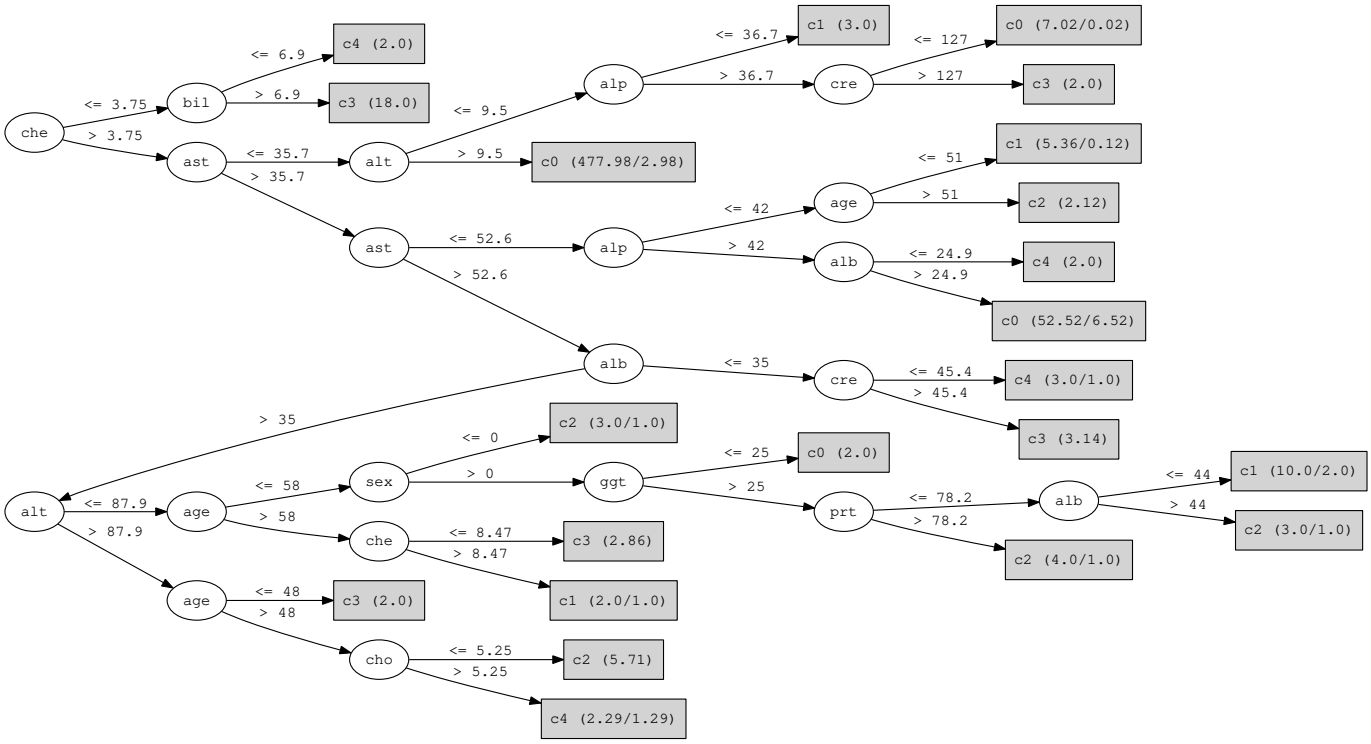


Fig. 11. Classifier generated by J48.

ACKNOWLEDGMENTS

This work was supported in part by the US National Science Foundation. It is based on a dataset published by prior work [1], and benefited from the reviewers' comments.

REFERENCES

[1] Georg Hoffmann, Andreas Bietenbeck, Ralf Lichtinghagen, and Frank Klawonn, "Using machine learning techniques to generate laboratory diagnostic pathways—a case study," *Journal of Laboratory and Precision Medicine*, vol. 3, no. 6, 2018. http://jlpm.amegroups.com/article/view/4401

[2] D. R. Hipp, D. Kennedy, and J. Mistachkin, "SQLite documentation," https://sqlite.org/docs.html, Jan. 2017, version 3.16.2.

[3] Eibe Frank, Mark A. Hall, and Ian H. Witten, "The WEKA workbench." 2016, online appendix for Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, Fourth Edition, 2016.

[4] William W. Cohen, "Fast effective rule induction," in *Proceedings of the Twelfth International Conference on Machine Learning*. Morgan Kaufmann, 1995, pp. 115–123.

[5] Jens Hühn and Eyke Hüllermeier, "FURIA: an algorithm for unordered fuzzy rule induction," *Data Mining and Knowledge Discovery*, vol. 19, no. 3, pp. 293–319, Dec 2009. https://doi.org/10.1007/s10618-009-0131-8

[6] Mark A. Hall, "Correlation-based feature selection for machine learning," Ph.D. dissertation, The University of Waikato, Hamilton, New Zealand, Apr. 1999.

[7] Abhishek Jha and Mrinal Sen, "Use of FES with vibrotactile feedback for motor learning," in *2016 International Conference on Internet of Things and Applications (IOTA)*, Jan. 2016, pp. 148–151.

[8] Kanika Sood, Boyana Norris, and Elizabeth Jessup, "Comparative performance modeling of parallel preconditioned krylov methods," in *2017 IEEE 19th International Conference on High Performance Computing and Communications; IEEE 15th International Conference on Smart City; IEEE 3rd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, Dec. 2017, pp. 26–33.

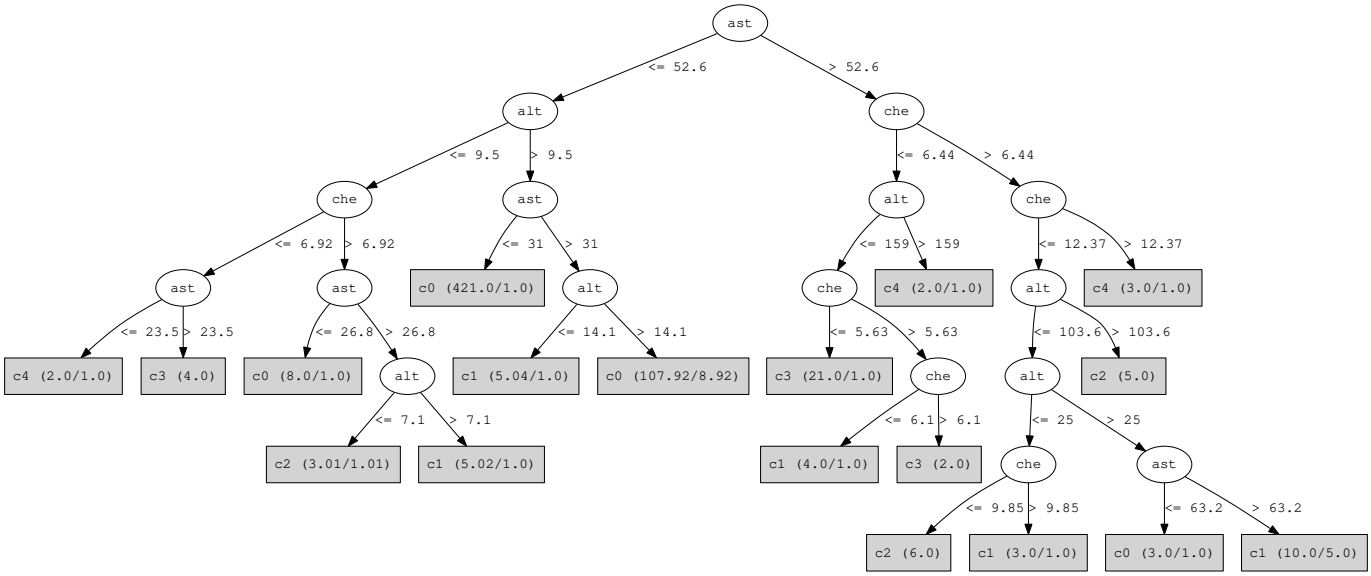


Fig. 12. Classifier generated by J48 using only three attributes: `alt`, `ast`, and `che`.

```
(ast in [-inf, -inf, 35.5, 37.8]) and (alt in [9.5, 9.8, inf, inf]) => cat=c0 (CF = 0.99)
(ast in [-inf, -inf, 52.6, 53.5]) and (alt in [10.2, 24.4, inf, inf]) => cat=c0 (CF = 0.98)
(ast in [-inf, -inf, 29.5, 31.1]) and (che in [3.75, 6.93, inf, inf]) and (alt in [3.8, 7, inf, inf])
    => cat=c0 (CF = 1.0)
(ast in [32.6, 37.8, inf, inf]) and (che in [9.45, 9.48, inf, inf]) and (alt in [-inf, -inf, 27.4, 27.5])
    => cat=c1 (CF = 0.58)
(ast in [35.7, 35.8, inf, inf]) and (alt in [103.6, 114, inf, inf]) => cat=c2 (CF = 0.51)
(ast in [36.8, 54, inf, inf]) and (che in [-inf, -inf, 4.37, 4.91]) => cat=c3 (CF = 0.86)
(alt in [-inf, -inf, 4.8, 5]) => cat=c3 (CF = 0.71)
(ast in [77.2, 80, inf, inf]) and (che in [-inf, -inf, 6.29, 6.65]) => cat=c3 (CF = 0.78)
```

Fig. 13. Unordered rules classifier generated by FURIA using only three attributes: `alt`, `ast`, and `che`.

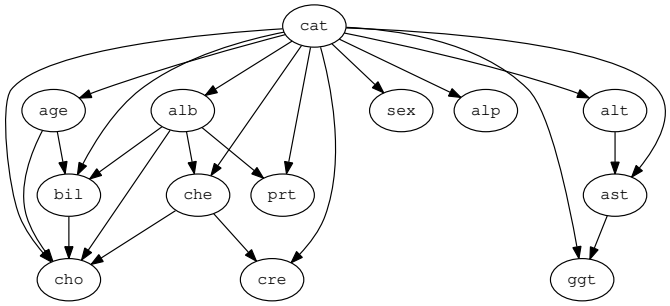


Fig. 14. Bayes Net for the 12-attribute classifier.

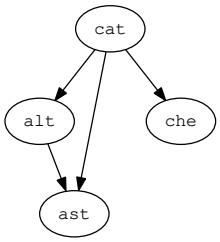


Fig. 15. Bayes Net for the three-attribute classifier.

[9] Archita Bahattacharya, Malay Gangopadhyay, and Sankhyabrata Bandyopadhyay, "Designing of highly sensitive fiber bragg grating based volume refractive index sensors for biochemical applications," in *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Nov. 2018, pp. 918–923.

[10] Reza Filsoof, Alison Bodine, Bob Gill, Stephen Makonin, and Robert Nicholson, "Transmitting patient vitals over a reliable zigbee mesh network," in *2014 IEEE Canada International Humanitarian Technology Conference - (IHTC)*, Jun. 2014, pp. 1–5.

[11] Srijita Chakraborty, Kaushik Das, Kaustavi Sen, Uttiyo Hari, Pritam Sarkar, Abhishek Raj, Malay Gangopadhyaya, Bob Gill, and Mrinmoy Chakraborty, "Elimination of higher order modes in wearable microstrip antenna," in *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Oct. 2016, pp. 1–4.

[12] Doina Bein and Bharat B. Madan, "Reducing the data communication delay in wireless sensor networks," in *2016 IEEE 12th International Conference on Intelligent Computer Communication and Processing (ICCP)*, Sep. 2016, pp. 361–368.

[13] Pratik Chattopadhyay, Shamik Sural, and Jayanta Mukherjee, "Information fusion from multiple cameras for gait-based re-identification and recognition," *IET Image Processing*, vol. 9, no. 11, pp. 969–976, 2015.

[14] Moises Levy and Jason O. Hallstrom, "A new approach to data center infrastructure monitoring and management (DCIMM)," in *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*, Jan. 2017, pp. 1–6.

[15] Per Bothner, "Kawa, the Java-based Scheme system," http://www.gnu.org/software/kawa/, Jun. 2003.

Implementation of an Informative Website – "Covid19 Predictor", Highlighting COVID-19 Pandemic Situation in India

Shuvankar Roy
Computer Science and Engineering Department
MCKV Institute of Engineering
Howrah, India
shuvankarroy2@gmail.com

Mahua Nandy Pal
Computer Science and Engineering Department
MCKV Institute of Engineering
Howrah, India
mahua.nandy@gmail.com

Sonali Bhattacharyya
Computer Science and Engineering Department
JIS College of Engineering
Nadia, India
sonali.bhattacharyya@jiscollege.ac.in

Srirup Lahiri
Computer Science and Engineering Department
JIS College of Engineering
Nadia, India
sriruplahiri1@gmail.com

***Abstract*—In this paper, we have represented implementation of an informative online platform- Covid-19 Predictor which is capable of disseminating accurate prediction of confirmed, deceased and affected Covid-19 cases in India on the basis of the data available in a reliable online repository. The work characterizes proper utilization of advanced technologies for web scrapping, model prediction, implementation of web application framework and cloud hosting.**

Keywords—covid-19, Web Scrapping, Deep Neural Network, Web Application Framework, Cloud Hosting

I. INTRODUCTION

The novel coronavirus disease 2019 (covid-19) pandemic created an urgent threat to global health causing disruption in daily livelihood. Despite government lock downs and public health responses aimed at containing the disease and delaying the spread, several countries have been crippled with a critical care crisis, and more countries could follow in future. Outbreaks lead to important increases in the demand for hospital beds and shortage of medical equipment, while medical staffs have higher probability of getting infected.

Future Projection plays an important role in mitigating the burden on the healthcare system, as well as providing the best possible care for patients, efficient diagnosis and information gathering on the prognosis of the disease. Prediction models that combine several variables or features to estimate the future numbers of affected due to the pandemic can provide required insights about it. It will also help policymakers to generate optimal policy for the upcoming situations. This will assist medical staffs to take necessary arrangements to allocate medical resources effectively.

Dynamic projection plays a key role in conditions like COVID-19 where, daily cases changes rapidly based on government restrictions, test conduction, infection rate etc. The contribution of the paper is as follows. In this paper, implementation of an online information sharing platform has been presented which exploits the technological advancements like web scrapping, prediction, web application framework, cloud hosting etc. to help common people to be aware of data analysis of COVID-19 infection in India. Our claim for dynamic system performance may be confirmed from, https://covid-19-predictor-india.herokuapp.com/ regularly.

II. LITERATURE SURVEY

In [1], R. Sujath et al. have presented a model that could be useful to predict the spread of COVID-2019. They have performed linear regression, Multilayer perceptron and Vector autoregression method for desire on the COVID-19 Kaggle data to anticipate the epidemiological example of the ailment and pace of COVID-2019 cases in India. They anticipated the potential patterns of COVID-19 effects in India depending on data gathered from Kaggle with the common data about confirmed, death and recovered cases across India for over the time length.

In [2], S. Lalmuanawma et al. addressed the recent studies that apply ML and AI technology towards augmenting the researchers on multiple angles. It also addresses a few errors and challenges while using such algorithms in real-world problems. [3] is an editorial report by A. Alimadadi et al. depicting Artificial Intelligence and Machine Learning methods applicable to Fight COVID-19. In [4], the N. S. Punn et al. propose to utilize the machine learning and deep learning models with the aim for understanding its everyday exponential behavior along with the prediction of future reachability of the COVID-2019 across the nations by utilizing the real-time information from the Johns Hopkins dashboard. In [5], Real-time data query is done and visualized in website, then the queried data is used for Susceptible-Exposed-Infectious-Recovered (SEIR) predictive modelling. Authors utilized SEIR modelling to forecast COVID-19 outbreak within and outside of China based on daily observations. Authors also claimed to analyse the queried news, and classify the news into negative and positive sentiments, to understand the influence of the news

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

to people's behaviour both politically and economically. M. Mandal et al. in [6], proposed a mathematical model to analyse COVID-19 pandemic both theoretically and numerically and evaluated the model by short term prediction in three states of India. V. Soukhovolsky et al. in [7], proposed mathematical model i.e. ADL (autoregressive distributed lag)-model in addition to the phase transition model to describe the development of the disease at the exponential phase. The ADL-model allows describing nonmonotonic changes in relative infection over time.

[8] and [9] represents two reports of WHO representing covid-19 pandemic situations at an early phase and at quite advanced phase respectively.

III. Dataset Description

The dataset [10] is contributed and updated on daily basis by Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE). This dataset is supported by ESRI Living Atlas Team and the Johns Hopkins University Applied Physics Lab (JHU APL). Johns Hopkins University is currently using a GitHub repository to store its data. It is in 2019 Novel Coronavirus COVID-19 (2019-nCoV) Data Repository.

In our implementation, we have utilized current country wise COVID-19 time series data of confirmed, death and recovery cases, available in [10] from 22nd January, 2020, to till date current data with an update frequency of once a day around 23:59 UTC.

IV. Implementation Methodology and Requirements

The implementation work has three different accomplishment phases, web scraping or web harvesting data, model implementation to predict the Covid-19 pandemic data and user friendly and visually influential online platform interface implementation to represent the data to the common people. Regularly the system automatically fetches updated data from the data source as and when the data source itself gets updated at their GitHub repository automatically and the models train themselves to predict next 5 days predictions without any human intervention. Figure 1 presents implementation methodology.

A. Web-Scrapping:

Web Scraping, also termed as Web Harvesting, is a technique employed to extract large amounts of necessary data from websites whereby the data is extracted and saved to a local database in structured format. In our implementation, the GitHub repository is being scrapped regularly for fetching *the* time series data of total confirmed, deceased and recovery COVID-19 cases in the Republic of India on a particular date. This repository contains day wise total confirmed, deceased and recovered cases of different countries i.e. daily time series summary tables of confirmed, deaths and recovered cases worldwide. In our local repository, this time series data, fetched from remote repository are transformed to the time series of cumulative growth of confirmed, deceased and recovered cases in India. To fine tune the model learning to capture the trend of daily change of values more precisely we modified the actual data format in our local repository.

Python Pandas is an effective tool to work with time series data. Website data fetching through remote URL path and data parsing from remote location to Pandas data frame directly form the backbone of the whole process. Indexing and extracting data from data frame enables data manipulation and transformation in an application specific way.

B. Deep Neural Network (DNN)

Artificial neural networks having multiple number of layers is called deep neural network. A fully connected (FC) neural network layer is said to be a dense layer. It accepts the input from each of the previous units and produces outputs for all the output units. Activation function helps in learning the patterns of characteristic of data. In this model Leaky ReLu has been used as nonlinear activation function which is capable of learning the nonlinear trend of data. It is defined as,

$$f(x) = max\ (a, x) \tag{1}$$

The architecture, present in Fig. 1, consists of only four layers of dense and Leaky ReLu. Total 116,161 trainable parameters and 0 non-trainable parameters. COVID-19 affected numbers and rates are completely dynamic in nature due to large population, government restrictions, number of testing conducted is variable per day. Therefore, projections needed to *be* corrected as per current data available for prediction of future with greater confidence. Our Deep Neural Network model is dynamically trained whenever the time series data is updated and the scenario of next five days is predicted. Confirmed case, deceased case and recovery cases are learnt by the models from regular *cumulative* growth values. Similar model architecture has been used for all of the predictions.

All the experiments have been implemented in Python 3.6.9. High level APIs of Google's deep learning library TensorFlow have been used to implement and evaluate the work. Implementations have been executed in Google online Colab cloud environment where Tesla P100 GPU, 16 GB VRAM, 1* Intel(R) Xeon(R) CPU @ 2.30GHz with 2 threads, 13GB RAM, 68.40 GB HDD.

C. COVID-19 Predictor Website Implementation

Website is based on client-server architecture concept and it is created using Django web-application framework. Website link is already provided for verification.

a) Web Application Framework: Django web application framework is preferred for implementation as it provides users certain components like database connectivity, authentication system etc and it also provides a lite weight server so that users do not need to install any third-party webserver. Following are the characteristics features of Django, which are very useful from developer's point of view.

- Django provides security to client-server data transmission by default.
- Django uses MVT(Model View Template) concept, model is for declaring database schema, view is for performing the operations and template is to show the output in client side.

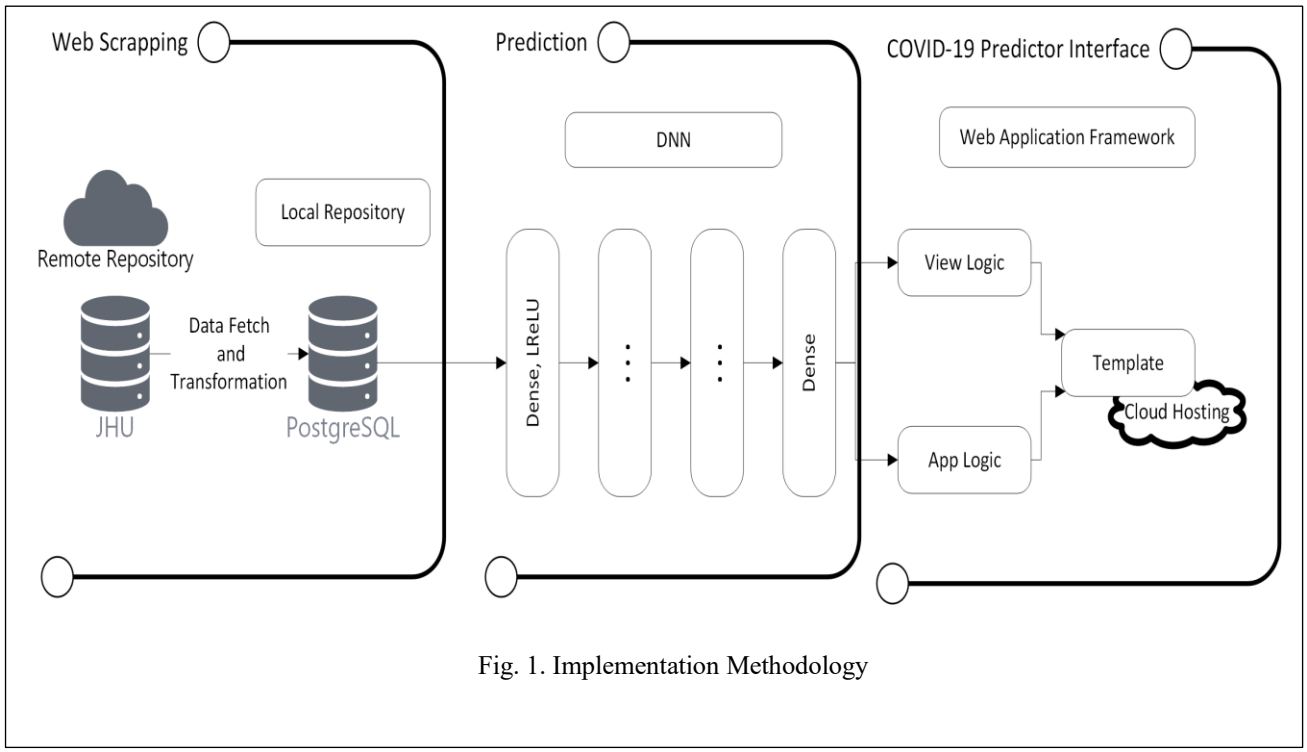


Fig. 1. Implementation Methodology

- Django use ORM (Object Relational Mapper) concept for database connectivity. In Django we need to create python class, ORM converts that class to a relation table which gets stored in the database. In this work, we have used PostgreSQL database and psycopg2 as on ORM.
- chartjs is used to plot the graphs using data fetched from server.
- Heroku cloud platform is used to host the server which is a platform as a service (PaaS), enables to build and run the application built on Django in cloud

b) Implementation Requirements: For server-side python programming language and Django module is used, to design client-side HTML, CSS and JavaScript is used.

Heroku dyno configuration for server-side operation:

- Operating System: Name of the OS is "Stack", which is based on Ubuntu (Linux based, open source operating system).
- The server has been deployed on Heroku on free dyno type with 2* Intel(R) Xeon(R) CPU E5-2670 v2 @ 2.50GHz, 512MB RAM.
- Python 3.6.10, tensorflow-2.0.0b1, Django-3.0.4, django-heroku-0.3.1 has been provided by Heroku to support the cloud hosting.

Any browser capable of rendering HTML5 with support of JavaScript and CSS for example Chrome, Internet Explorer 10+, Microsoft Edge and Safari along with their mobile counterparts will be able to render the website at the client end. At our end, we have used Google chrome version 81.0.4044.129(official built) (64 bit) for our implementation.

V. Experimental Results

Covid-19 Predictor disseminates deep learning-based prediction of confirmed cases, recovery cases and death cases of Covid-19 pandemic in India.

A. Results Prediction and Informative Website

Three different prediction results of Covid-19 Predictor are represented in tabular form in tables I, II and III respectively. High values of accuracy indicate the effectiveness of the system. TABLE I, II and III represents Error and Accuracy Computation of Total Confirmed Case, Total Deceased Case and Total Recovery Case respectively. Figure 2, 3, 4 and 5 displays significant website interfaces. Prediction performance and the representations may be verified from the link provided, https://covid-19-predictor-india.herokuapp.com/.

TABLE I. Accuracy Computation of Total Confirmed Case Prediction

Date	Total Confirmed Case Prediction and Accuracy		
	Actual Confirmed Case	*Predicted Confirmed Case*	*Percentage Accuracy*
2nd August, 2020	1803695	1817160	99.25900856
3rd August, 2020	1855745	1871579	99.1539764
4th August, 2020	1908254	1924992	99.13048989
5th August, 2020	1964536	1978740	99.28216946
6th August, 2020	2027074	2036580	99.53323709

TABLE II. ACCURACY COMPUTATION OF TOTAL DECEASED CASE PREDICTION

Date	Total Deceased Case Prediction and Accuracy		
	Actual Deceased Case	*Predicted Deceased Case*	*Percentage Accuracy*
2nd August, 2020	38135	38408	99.2841222
3rd August, 2020	38938	39198	99.33227182
4th August, 2020	39795	40015	99.44716673
5th August, 2020	40699	40893	99.52332981
6th August, 2020	41585	41819	99.4372971

TABLE III. ACCURACY COMPUTATION OF TOTAL RECOVERY CASE PREDICTION

Date	Total Recovery Case Prediction and Accuracy		
	Actual Recovery Case	*Predicted Recovery Case*	*Percentage Accuracy*
2nd August, 2020	1186203	1203808	98.51585268
3rd August, 2020	1230509	1245616	98.77229667
4th August, 2020	1282215	1292083	99.23039428
5th August, 2020	1328336	1347921	98.52559894
6th August, 2020	1378105	1395666	98.72571393

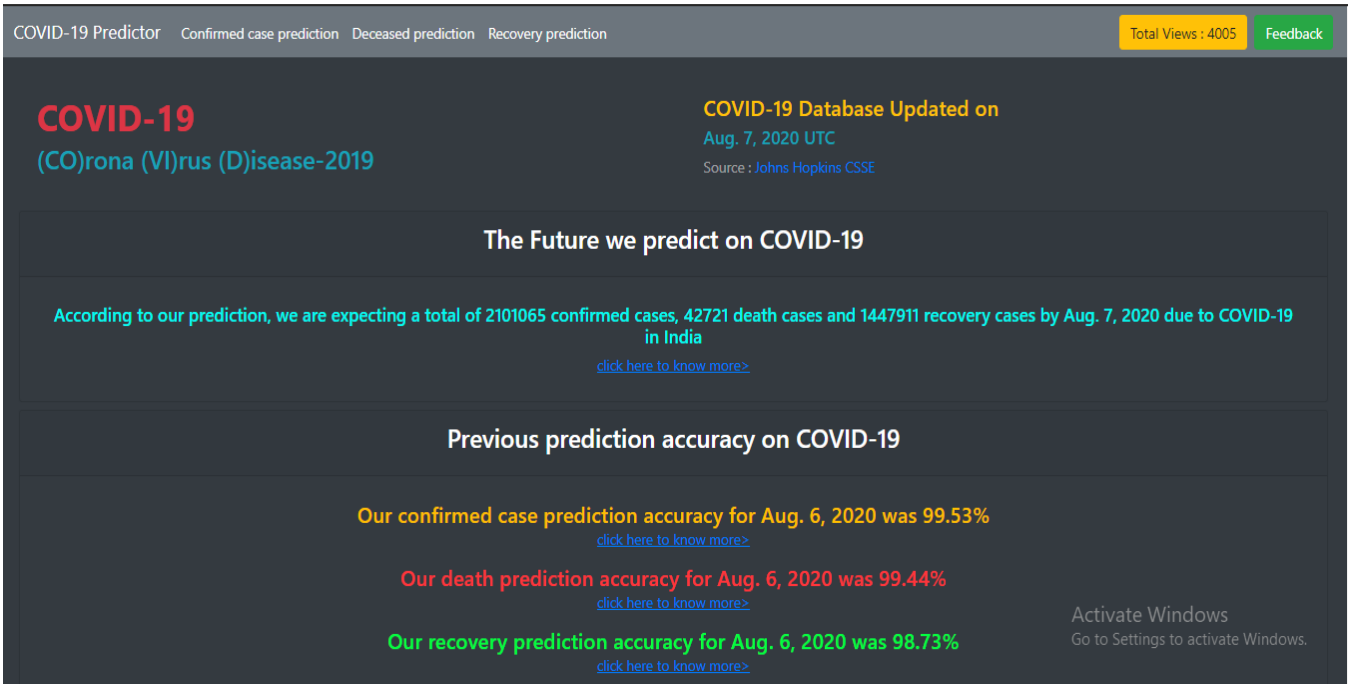


Fig. 2. Covid-19 Predictor Interface

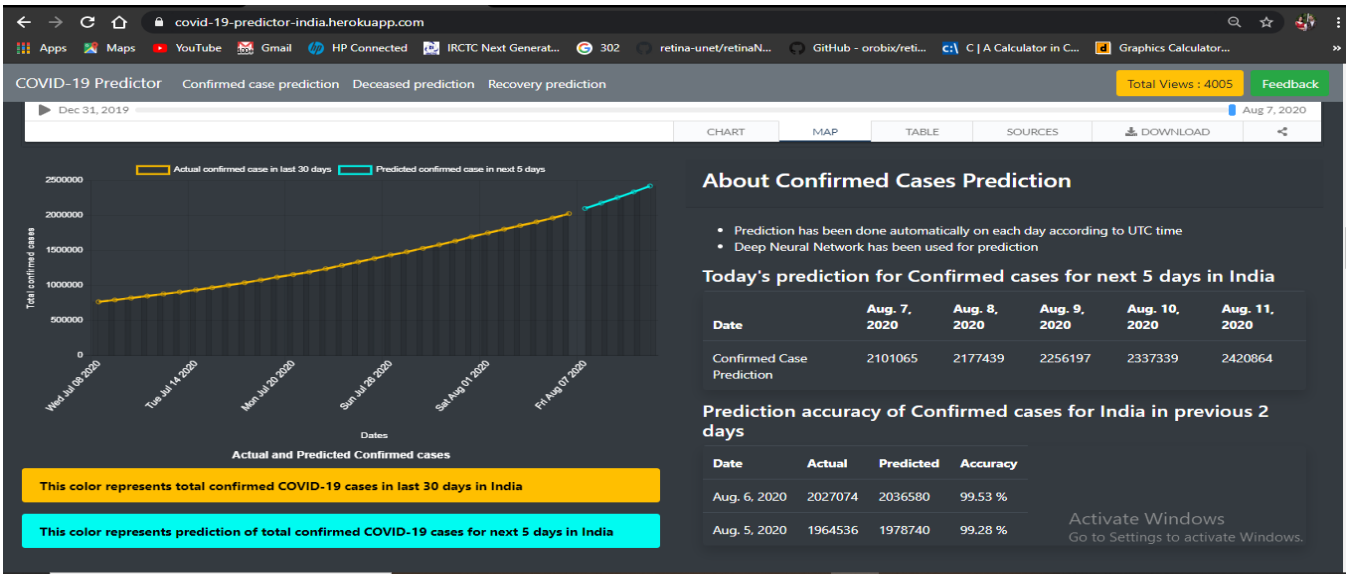


Fig. 3. Online visualization and prediction of Confirmed Cases

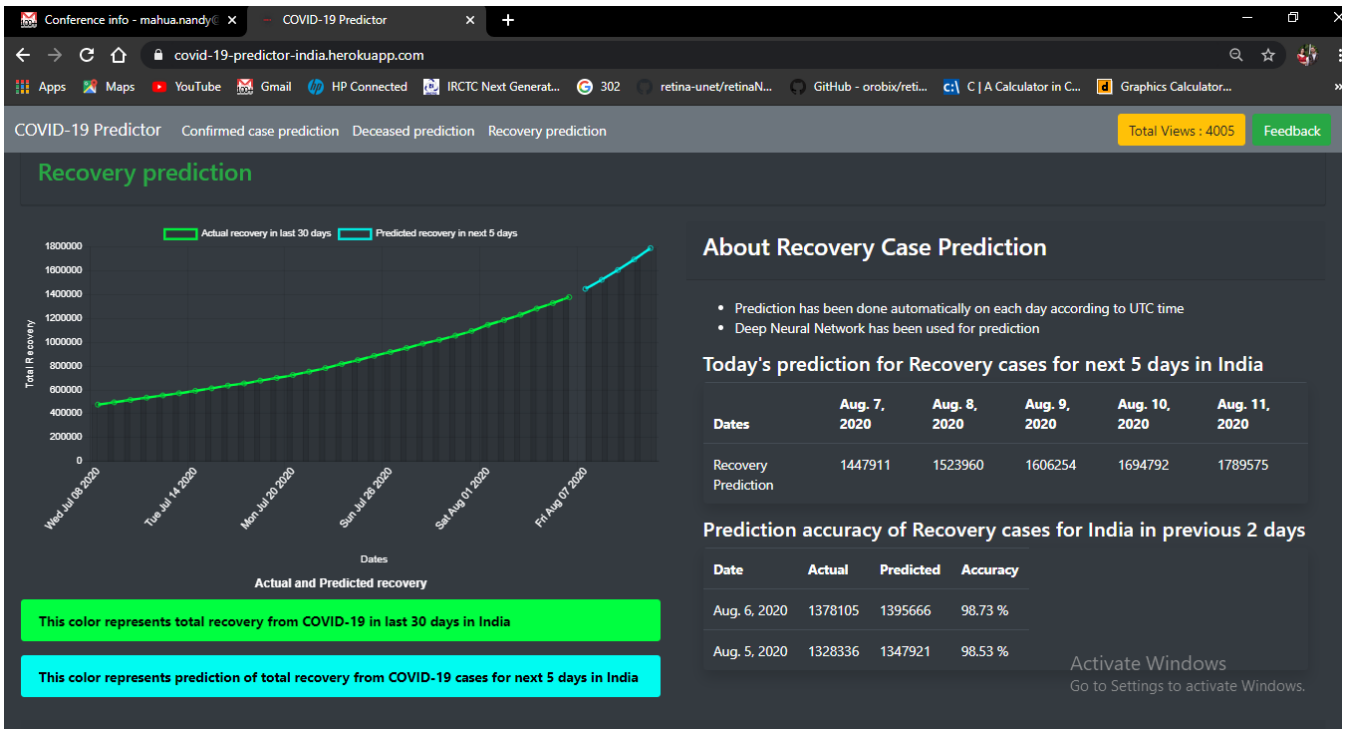


Fig. 4. Online visualization and prediction of Recovery Cases

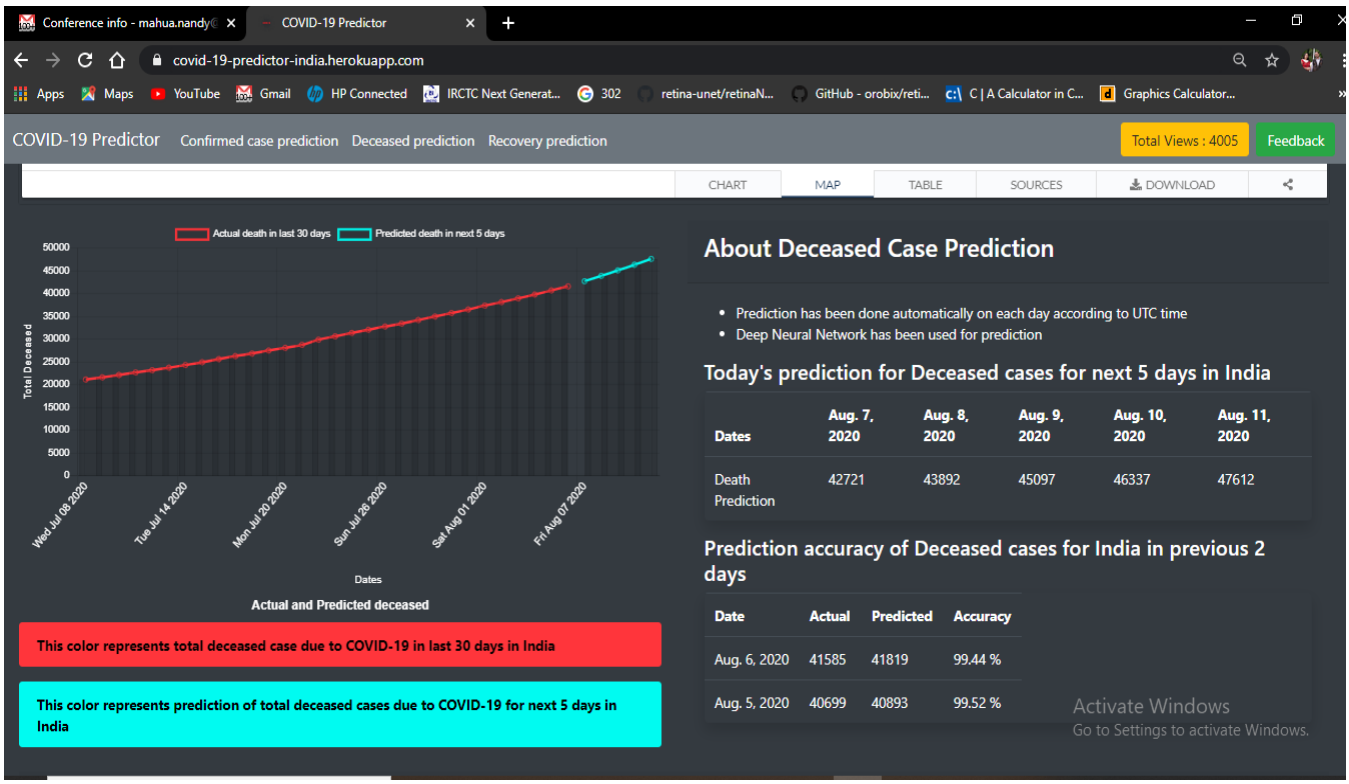


Fig. 5. Online visualization and prediction of Deceased Cases

CONCLUSION

Information dissemination to the people allows them to have proper visualization power about the spreading of the disease which helps them in establishing their own opinions during pandemic situation. Thus, they become aware of the requirements of different containment measures also such as travel ban, border closing, trade control etc. as well as sanitization measures as per WHO's directives. and provide support to the implementation authority spontaneously. This in turn, provides more successful imposition of human beneficiary actions during pandemic.

REFERENCES

[1] R. Sujatha, J. M. Chatterjee, and A. E. Hassanien, "Correction to: A machine learning forecasting model for COVID-19 pandemic in India," Stochastic Environmental Research and Risk Assessment, vol. 34, p. 959-972, 2020.

[2] S. Lalmuanawma, J. Hussain, and L. Chhakchhuak, "Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review," Chaos, Solitons & Fractals, vol. 139, p. 110059, 2020.

[3] A. Alimadadi, S. Aryal, I. Manandhar, P. B. Munroe, B. Joe, and X. Cheng. "Artificial intelligence and machine learning to fight COVID-19," Physiological Genomics, vol. 52, p. 200-202, 2020.

[4] N. S. Punn, S. K. Sonbhadra, and S. Agarwal, "COVID-19 Epidemic Analysis using Machine Learning and Deep Learning Algorithms," 2020.

[5] F. A. B. Hamzah, C. H. Lau, H. Nazri, D. V. Ligot, G. Lee, C. L. Tan, M. K. B. M. Shaib, U. H. B. Zaidon, A. B. Abdullah, M. H. Chung, C. H. Ong, and P. Y. Chew, "CoronaTracker: World-wide COVID-19 Outbreak Data Analysis and Prediction," Bull World Health Organ, 2020.

[6] M. Mandal, S. Jana, S. K. Nandi, A. Khatua, S. Adak, and T. Kar, "A model based study on the dynamics of COVID-19: Prediction and control," Chaos, Solitons & Fractals, vol. 136, p. 109889, 2020.

[7] V. Soukhovolsky, A. Kovalev, A.Pitt and B. Kessel, "A new modelling of the COVID 19 pandemic," Chaos, Solitons & Fractals, p.110039, 2020

[8] World Health Organization (WHO), "Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCov)," WHO, 2020.

[9] World Health Organization (WHO), "Coronavirus disease 2019 (COVID-19) Situation Report- 13," World Health Organization, 2020.].

[10] Dataset link: https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series

Stock Market Prediction using Text-based Machine Learning

Tristan Jordan
School of Engineering and Computer Science
Morehead State University
Morehead, KY, USA
tajordan2@moreheadstate.edu

Heba Elgazzar
School of Engineering and Computer Science
Morehead State University
Morehead, KY, USA
h.elgazzar@moreheadstate.edu

***Abstract*— Predicting stock market price movements can be a difficult task for traditional algorithms as random events can vastly change a stock's value. The goal of this research project is to design machine learning algorithms to predict these changes based upon communal discussion. The discussions that are being analyzed will be from forum posts of users that have varying levels of involvement with the company of focus. The posts themselves should contain information related to the current events, problems, community sentiment and other factors that would influence buyers and sellers. The proposed algorithm to make these predictions with a recurrent neural network (RNN) that will be able to analyze patterns in word use and order, placing reactions to forum posts into a category based upon expected price movement over various lengths of time. These methods show promise in predicting performance over many time frames.**

***Keywords*—Stock Market, Recurrent Neural Network, Text based predictions**

I. Introduction

The stock market is a dangerous beast with many factors leading to the seemingly unpredictable gains and losses. Many of the traditional ways of making the market predictable included performing traditional algorithms on the numerical data. These regression models have good accuracies, however, react slow to major shifts in public events.

With the amount of data that is poured into the internet at any given moment it is important to be able to distinguish meaning. Extracting a meaning from text is an idea known as text mining providing many uses within many problems [1]. In the scope of this project text mining is used to understand sentiment within forum posts. The hope of this project is that people who are invested within a company would take in information related to the company then sharing their opinions on public messaging boards. These opinions are then analyzed for their thoughts on a topic or various other implicit attributes such as passion, authority, and intelligence to produce a guess about movement in a company's stock.

This project is two parts: creating the dataset and learning from the dataset. This is done in the absence of a proper dataset for stock market predictions using forum posts with the price "reactions" after various intervals.

The predictions on this data is a class that says the level of increase (+-10%, +-5%, 0%). Isolating these values allows an efficiency to investing to a respective stock to allow a user to maximize gains during a specific time. Time frames are split up to get an idea how the predictions scale. The algorithm tries to pull data as close to the date as possible, it does not always retrieve data for the exact date due to restrictions from the API. The time frames that this project uses are one day, one month, three months, one year, and five years.

This paper discusses the methods of a stock market predictor using machine learning. The paper is organized such that: Section 2 explains similar research that has happened within the field. Section 3 details the steps taken to construct the algorithm furthered by the inner workings of the choices made. The results are the focus of section 4. Ultimately, the conclusion is provided in section 5.

II. Related Work

Research into predicting stock market fluctuation is a lucrative investment for many individuals. This leads to many of the systems used to be a black box with much of the research to be hidden.

Green Key Technologies performs stock market analysis using many features from their other software [2]. These features allow them to classify different reports using Natural Language Processing. This analysis is provided to the user within their platform that also identifies the user's past conversational data.

In 2018 Queen's university attempted predicting a stock's performance using articles provided by moneycontrol.com [3]. Their method was using a dictionary to account for words that would be used within the article then develop a sentiment value for each key, these sentiments were then summed up to get a total sentiment for the article. Their predictions were "buy", "hold" and "sell"; using these methods they saw an accuracy of 70.59% showing promise for using text to predict financial decisions in the stock market. Issues mentioned with their approach were mainly focused on most of the work with the dictionary being manual.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Desheng Dash Wu and David L. Olson saw success in 2014 using analysis on a stock market forum from china. Their method was to perform sentiment analysis on a post then run that data through an SVM (Support Vector Machine). They saw an accuracy of 81.82%, during their testing they found that more posts caused a better accuracy.

The need the ever-flowing stream of data is ever increasing. Using machine learning methods are necessary to be able to understand information at the rate it is being created. The related work shows that there are enough relationships within text data to be able to reliably draw conclusions.

III. Proposed Method

As previously mentioned, this project consisted of two parts, being a web crawler and price predictor. Data preprocessing occurs during both parts with different objectives. Though being different pieces, the project shared many of the same tools.

The programming language chosen for this project was Python 3. This language is a data scientist's paradise due to its simple syntax combined with well written libraries. Anaconda is the platform that was used to manage packages, allowing for simplified state management. HTML parser and URL lib were essential to the crawling operations. Additionally, NumPy and Pandas were essential for various operations including reading files, writing files, or managing data [8][11]. The neural network was written using TensorFlow with the Keras framework allowing for easy development [9][10].

A. Dataset

The goal of creating the dataset is generating an easily accessible and low-cost set of documents which will require little further alteration. The forum posts for the project is mined from Investors Hub, an advanced stock discussion platform. This site hosts over 20 years of forum posts with a diverse set of public users, boasting ~40,000 posts per day across over 30,000 boards. The stock performance is taken from the Alpha vantage API, with a history matching Investors Hub. These sources of data will populate CSV files with their information.

1. Web Crawler

Web crawling is a lengthy process in which a computer program automatically pulls information from a site. The crawler mentioned is pulling forum posts, dates, and company stock tickers from Investors Hub to be saved for further analysis. This is a lengthy process that requires many items to be well thought out before starting the crawl.

The first consideration to be made when crawling a site is to do so in a way that is ethical. Being ethical in crawling refers to limiting requests to prevent giving too much traffic to the host. Failing to crawl ethically can lead to the connection being blacklisted, stopping all crawling. Achieving an ethical crawl was done simply in this project, each time a post downloaded the program would sleep for half a second before resuming. Additionally, if there was a failed download the program would wait 60 seconds before moving on to the next item to be parsed. the next consideration that must be performed is what data is being searched for.

Web crawling is best done when an idea is presented, then the machine follows the algorithm to get the data. An example of an item that the program may search for:

"data"

The first step to data mining is retrieving the html text of a website. If one were to search for a link like this, they would first look for an the "a" tag that usually contains links. They would then try to search for the word "goal" in the id attribute. When a desired item is found the data is extracted then saved to its respective file. To find the instructions a developer console was used, and the site used clear naming schema on the elements within the site. A recursive mining hierarchy is shown in Fig. 1.

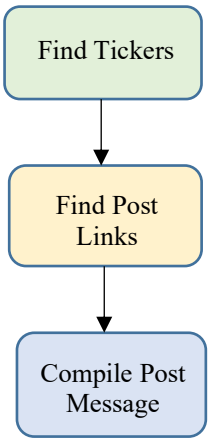


Fig. 1. Recursive Mining Hierarchy, more iterations performed closest to the bottom

The first item to be mined off the website are the tickers for each company. A ticker is a collection of letters that represents the company, for example, GOOGL is the ticker for Alphabet Inc. In the project the ticker is used to label the respective dataset, also, leads the request to pull data from the Alpha vantage API and the respective stock page. In this step the tickers are pulled into an array, shuffled, then pushed into a NumPy array to be saved. Saving the tickers allows for the crawl to be paused when needed.

During the main crawl, the ticker file is loaded into memory to be fed one by one into the web scraper. A set number of posts from each stock is downloaded, formatted then saved. The set number is to try to keep every ticker with roughly the same amount of posts as to not bias towards certain companies.

One by one links mined from the ticker's forum page are parsed. The parse looks for all inner HTML from body div. This process also includes keeping track of passed tags as to ignore any data that is not related to the forum post itself.

One API call is made to Alpha Vantage per stock ticker. The information sent includes function (daily stock prices), ticker symbol, datatype, and the API key. Alpha vantage then returns a JSON which contains price information over the company's lifetime. A page is then searched over for each stock on the Investors Hub page which lists links to all posts that exist for each entry. A selected group is pulled out with the post along with the date. This data allows 6 different dates (original date, one day, one month, three months, one year, and five years) to be calculated, each being matched with the highest price of a given day taken from the JSON. These dates returned from the API are not always included for various reasons (weekends, holidays, etc.). If the exact searched date is not present, the day closest to it (while being before the expected date) is selected. If the selection falls on or before the date of the forum post, then the fluctuation for that period is replaced with "NA".

Prior to being saved the forum post undergoes the following transformations:

- Removing non-ASCII characters
- Removing HTML escapes
- Removing line breaks
- Allowing only one copy of each character per word
- Lowercasing every character
- Removing extra space
- Removing special characters

Many of these operations were done to reduce the amount of unique words to provide space in the dictionary along with reducing the size of the saved data overall. The dictionary is the way that the program can transform text data into arrays of discrete numbers. It does so by having a blob of text, finding the top X words used within the blob then assigning each *X* a unique key within the dictionary (if the dictionary is full throw *X* out).

Completing the first wave of preprocessing the data is then transferred to a NumPy array and saved. The structure of the file that is saved is a csv titled by the ticker structured as:

Message	*Date[0]*	*Price[0]*	*Date[N]*	*Price[N]*

With *N* being the index at a date marker (*day*, *month*, *year*, *etc.*).

2. Predictor

Predicting the data should be accurate as well as fast as the goal of the project is to keep up with the demand of data analysis without human intervention. This model should be able to guess a range of movement, with one model per date range. The ranges that the model will try to predict are +-10%, +-5%, and 0%. These ranges should tell a user the chances that money will be made from a stock, leading to how much money one would expect.

Language is often difficult to interpret for a machine. A Recurrent Neural Network (RNN) is employed so that context can be pulled from the incoming data. The context will be trained on similar data to that of the focus discourse platforms to ensure expressed mannerisms have been seen before. The network will learn like a traditional network, with guesses being made followed by adjustments to the neurons through backpropagation to minimize the loss (categorical cross entropy).

a) Preprocessing Data

Prior to running data through the machine learning model, it is important to preprocess all the data. During this step of preprocessing there is a major focus on the output data and general discretization of data. It is important that data is discrete to ensure that enough of the data is recognizable to the model.

The first step of preprocessing at this level involves tokenizing the input. During this step words are changed to codes that are predetermined in the dictionary. A maximum of 50,000 unique words are stored into the dictionary with favor given to frequency. Any entry that cannot be fit into the dictionary is thrown away as its scarcity likely will be consistent in future tests. All the posts will be arrays of integer codes at this point.

The prediction from the model will be observed through:

$$\%\ change = \frac{Price2 - PriceOrigin}{PriceOrigin} \qquad (1)$$

The new value is represented as a percent change. These percentages are separated into buckets to discretize the value and allow encoding. The process of one hot encoding entails each value being an array with a length equal to the largest integer in the list every value being 0 except for the index equal to the integer that the change corresponds to. For example:

$$OneHotEncode(3) = [0,0,0,1,0]$$

At the index of 3 the value is 1 given a max length of 5, allowing the machine learning to compute a percent chance that each index is correct, with the largest value being the overall prediction.

b) Training the model

Neural Networks are employed to create predictions from large amounts of data in a non-linear format. This subset of machine learning takes many nodes or neurons and layers them on each other creating a deep schema (deep learning) [5]. This method of machine learning was chosen for the high dimensionality contained within the dictionary for each post.

Keras makes creating these neural networks simple as you build your model around the concept of layers. Each layer has its own attributes as the data flows into a layer from the previous and out of the layer into the next or an output layer. There are some exceptions to this rule such as recurrent layers. These special layers will retain partial information to preserve context of incoming data. The model used in this project has three primary layers as shown in Fig. 2. This model should not be trivial as it is run roughly once per word in a post like a chain.

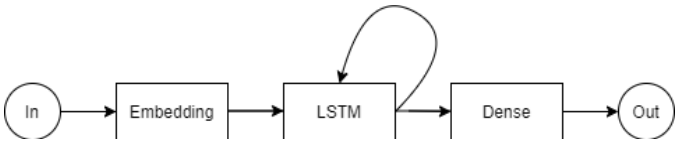


Fig. 2. Price Predictor Model

The input into this model is an array with a predefined length. The way this length is found is the longest forum post in the training set. For entries with too few words, nulls pad out the extra space. Arrays with too many words will trim the end changing its length into the max size.

Each neuron in a layer takes in a sum of all inputs from the output of the previous layer. This sum is then applied with an activation function which dictates how a neuron "fires" giving the model its nonlinear attributes. Activation functions often limit the output that a neuron can produce within the network.

The activation function that is used in this work is known as a "Sigmoid" [6]. This activation function is nonlinear, always outputting a value between 0 and 1. The sigmoid function can be expressed mathematically as:

$$f(x) = \frac{1}{1+e^{-x}} \qquad (2)$$

The first layer in this network is an embedding layer. These layers will draw similarities between words providing them interchangeability using a similarity vector. The vector can be imagined as a graph with N dimensions (definable within the layer), words that are similar will be close to each other. The primary purpose of this layer is to generalize the meanings of words and allow for more insight into posts that the network may not see during training. The output of this layer is the similarity vector previously mentioned.

The second layer of this network is a Long Short-Term Memory (LSTM) layer which allows the model to become an RNN [7]. These layers hold pieces of information from runs that are in succession which is usually lost between runs. These pieces of information give contextual based learning to the network where certain words may change their meaning entirely based upon their position in a sentence. An example of this would be "not" and "good", when used as "the outlook is not good" the phrase is synonymous with "bad outlook" as opposed to a phrase possibly saying "good not a bad day yet" which furthers the thought of the performance continuing to be good.

A drop out of 5% is applied to the LSTM layer. This feature allows for random connections to be turned off before the output to the next layer. These cutoffs help to force the layer to make connections based upon all data in the layer as opposed to key points which, in rare cases, may be omitted from the input. Additionally, a recurrent drop out of 5% was added which blocks partial information from the next run of the same level. Fig. 3 shows an LSTM cell that carries data between waves of inputs.

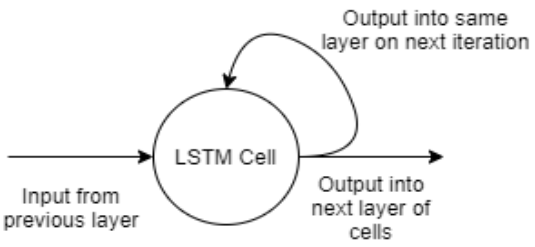


Fig. 3. LSTM Cell preserving data between epochs that would normally be lost

The final layer in the model is a dense layer. These layers are often seen as the traditional neural network layer. They map input nodes to output nodes and perform their activation function to find implicit connection of the input data to the output data. This layer is important as it provides many of the traditional learning sought from the model. An example of fully dense neural network is shown in Fig. 4. The output for this model is a list of values from 0-1 with the same length as the one hot encoded input. Each value is the model's prediction that the text belongs within that class in percentage. Pictured below is a simple neural network with a simple layering known as "dense":

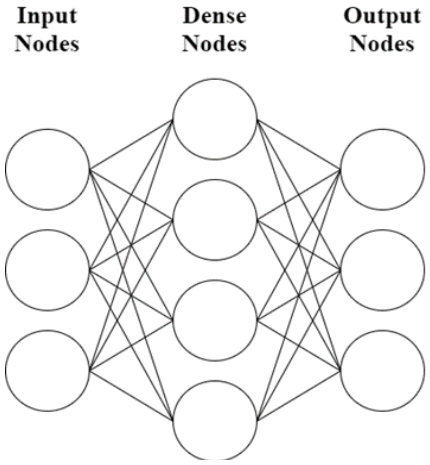


Fig. 4. Example Fully Connected Dense Neural Network

c) Saving the model

The final step is to save all data that may need to be used in future runs. The program stores the training and test data used during runs along with the model's files. The file that the model saves to is ReccurrentClassifier.h5 like a pickle, as reloading the model is performed primarily by the library.

IV. RESULTS

Experiments were conducted to test the accuracy of the proposed algorithm. The experimental results for the algorithm were calculated in two ways, the first being the ability for the algorithm to guess the exact class correctly (the exact class 1-5)

with the second being the algorithm predicting the trajectory correctly (gains or losses). Roughly 33,000 forum posts were collected with a split of 90% used for training alongside 10% testing. The network trained over 50 epochs (waves) per run with the loss function being categorical cross entropy. The prediction accuracy results are shown below in Table I.

TABLE I. PREDICTION ACCURACY RESULTS

	Predicting the Class	**Predicting the Trajectory**
1 Day	76.64%	86.97%
1 Month	71.48%	82.56%
3 Month	63.87%	76.07%

It is important to note that this accuracy is on a per post basis. Commonly on stock discussion boards people will hold onto an opinion with little in the way of convincing them away from it. A way to combine or average out guesses for posts over a period may see better results. A post showing that a stock will perform well may also have a higher chance of being an increase stock overall as well, so that someone who was using the algorithm to make predictions would solely target that class.

The algorithm also returns a list of all class probabilities, so performing different analysis on those values could also prove to be valuable. Additionally, there may be people on the discussion board who are intentionally providing false information that would have slipped passed the site's security. Another interesting pattern with the data is that longer periods of time produce better results. This could be due to a wide variety of factors including reverse stock splits (combining shares of a stock to raise the price) or general inflation which have not been filtered out.

Ways that could be seen to optimize the accuracies presented in this model may include seeking out "dummy posts" which add no value to the discussion. These posts are often from users that hold either strong or no feelings towards the company. Similarly, if there were discussions occurring, it would be beneficial to stop the program from "double dipping" from certain users' interactions, such as repeating the original post or going off topic. Additionally, a way of bundling post scores over periods of time could see success.

V. CONCLUSION

The employment of computer application to predict trade values can be useful for the modern broker. Many algorithms have been created to perform analysis that see success from their ventures. Most lack the verbal discussion that happen from those involved. This paper tries to investigate the viability of investor discussion which can sometimes predict important action before numerical data.

The need to collect data comes from the highly volatile nature of the stock market. Data will become stale quickly, only allowing the model to learn from the old information will miss many important pieces which could give a much better output. The methods used for data mining would be able to keep the training data up to standard with any evolutions that may take place.

Neural networks were utilized as their nonlinear structures allow for many implicit interactions to be preserved during analysis. Moving away from the traditional feed forward network into a recurrent neural network, the ability to understand the posts within context is added. This context proves to assist in producing reliable output.

Machine learning in the context of stock market predictions shows promise as a valid member in an investor's toolkit. The experimental results show that there is a correlation between user posts with their corresponding movement.

The proposed algorithm still has many additions together with modifications that could be made to increase its effectiveness. The first thing that needs to be done is downloading more data as well as more experiments with other forms of preprocessing. A way of being able to interject new movement in language would also assist in the rapidly changing nature of modern language. Additionally, better analysis of better ways to use the output from the neural network to make better decisions could see large differences (such as the class accuracy against the trajectory accuracy). Finally, more simulation with decisions that would be made directly from the data would give a better idea of how integrating these tools into investment would be.

REFERENCES

[1] M. Allahyari, S. Pouriyeh, M. Assefi, S. Safaei, E. D. Trippe, J. B. Gutierrez, K. Kochut, "A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques" arXiv: 1707.02919 pp. 1 (2017)

[2] "Natural Language Processing for Financial Markets - Convert Your Customer Conversations into Revenue Producing Insights," GK, 21-Apr-2020.

[3] D. Shah, H. Isah, F. Zulkernine, "Predicting the Effects of News Sentiments on the Stock Market" pp.1-4 (2018)

[4] D. Wu, L. Zheng, D. Olsen, "Online Stock Forum Sentiment Analysis" pp. 1077-1087 (2015)

[5] R. Saritha, V. Paul, P. Kumar, "Content based image retrieval using deep learning process" Cluster Computing, pp. 4187–4200 (2018)

[6] C. Nwankpa, W. Ijomah, A. Gachagan, S. Marshall, "Activation Functions: Comparison of trends in Practice and Research for Deep Learning" pp.1-20 (2018)

[7] A. Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network." pp.1-43 (2020)

[8] S. Walt, S. Colbert, G. Varoquaux, "The NumPy Array: A Structure for Efficient Numerical Computation, Computing in Science & Engineering", pp. 22-30 (2011)

[9] F. Chollet et al. "Keras" (2015). Software available from keras.io

[10] M. Abadi et al. "TensorFlow: Large-scale machine learning on heterogeneous systems", (2015). Software available from tensorflow.org.

[11] Pandas Development Team, "Pandas", (2020). Zenodo.

A Corpus of BTC Tweets in the Era of COVID-19

Toni Pano, Rasha Kashef
Electrical, Computer, and Biomedical Engineering Department
Ryerson University
Toronto, ON, Canada
toni.pano@ryerson.ca, rkashef@ryerson.ca

Abstract: The coronavirus disease 2019 (COVID-19) outbreak has created a great challenge for many industries, including healthcare systems, E-commerce, and transportation. Cryptocurrencies, such as Bitcoin (BTC), is an alternative class of digital assets primarily used as a medium of exchange. In the financial market, data collected from social media such as tweets, posts, and can assist in an early indication of market sentiment. This has frequently been conducted on Twitter data. In particular, this paper provides a corpus of tweet text for Bitcoin-related tweets during the summer of the COVID-19 era. This dataset is publicly available covering a considerable period of roughly three months, to allow people with a similar interest in Twitter, Bitcoin, sentiment analysis, and the financial sector to perform research unimpeded.

.

Keywords—Bitcoin, Sentiment Analysis, Tweets, Covid-19

I. INTRODUCTION

Bitcoin is an electronic cash system that does not depend on financial institutions to keep track of its payments. Twitter Sentiment analysis is the automated process of analyzing tweets data and sorting it into positive, negative, or neutral sentiments. Using sentiment analysis to analyze opinions in Twitter data can generally help companies to get a better insight into customers' behaviors and interests.

As observed in [1-6], tweets' sentiments on BTC could serve as an indicator of Bitcoin prices. The collected corpus of the BTC tweets during the COVID-19 situation could predict the BTC trends and their correlation to its sentiment. Any observations of association may be of interest to the financial sector for price prediction or risk mitigation. The dataset is useful for forecasting the behavior of Bitcoin during pandemics, similar to COVID-19.

Access to specific tweets is usually restricted behind a paywall or by Twitter. Also, any publicly available tweets are limited to a historical depth of 7 days by Twitter's Search API [1] or are too general. Thus, this paper introduces the covid-19 BTC tweets dataset as a useful corpus for many research fields related to text mining, price prediction, financial analytics, and trend prediction.

Sentiments extracted from Twitter using Natural Language analysis may correspond with the behavior of Bitcoin prices. This is consistent with the observations of multiple works of research done by E. Stenqvist and J. Lönnö [1], O. Kraaijeveld and J. D. Smedt [2], T. R. Li et al. [3], S. Mohapatra, N. Ahmed, and P. Alencar [4], C. Kaplan, C. Aslan, and A. Bulbul [5], and A. Jain et al. [6]. The sentiments expressed in the dataset were intended to be correlated against the price of Bitcoin during COVID-19, specifically for observing any changes in correlation. Groups of research interest include bankers, investors, or other parts of the financial sector that deals with trend prediction or risk estimation. The dataset may be helpful to anyone interested in sentiment analysis or other natural language processing related to Bitcoin. The dataset is currently employed in a project for sentiment analysis and machine learning.

The rest of this paper is organized as follows: in section 2, the data description is provided. Section 3 summarizes the data specification. Experimental design and methods are discussed in section 4. Finally, conclusion and future directions are presented in section 5.

II. DATA DESCRIPTION

Tweets related to Bitcoin were collected by accessing Twitter's free REST APIs [7]. Specifically, the Search API was used to search for tweets created in the past [8,9]. A tweet scraper was built to collect the data and store it in a CSV file format. All relevant information for processing tweets on a timescale was included if it did not break Twitter's Developer Agreement [10]. In the interest of guarding user privacy, the username of each tweet's account was not recorded, and further preprocessing plans for the removal of usernames in text. Each tweet posted on a specific day and month has been stored in a file beginning with "month-day-2020". As the index of each row in a file increases, the timestamps of the recorded tweets decrease, starting around 23:59:59 and ending at 00:00:00.

III. DATA SPECIFICATION

All data in each file has been stored as a string, and certain data columns that don't apply are left blank. Each row of the dataset contains all related info for a single tweet, identified by the "id" column. The "time", "id" and "text" of any requested tweet comprises the first three columns, as a result of this labeled as the "requested tweet." The "time" column stores the timestamp of the tweet creation as "YYYY-MM-DD hh:mm:ss" in a 24-hour format. The "id" column stores the unique 19-digit number of each requested tweet.

978-1-7281-9615-2/20/$31.00 ©2020 Crown

The "text" column stores the text of the requested tweet. If the requested tweet was a retweet of someone else's tweet, then the next three columns are filled with the "original time", "original id", and "original text" of someone else's original tweet. In the case of the requested tweet being a retweet, the last two columns are filled by someone else's tweet info. If there requested tweet was not a retweet, then the last two columns are filled by the requested tweet. If the respective tweet text happens to be greater than 140 characters, the untruncated text fills the "full text" column with up to 280 characters [11]. If the respective tweet text happens to quote another tweet, a quotes object is formed and stored in a Javascript Object Notation (JSON) format, under the "quotes JSON" column. The quoted tweet is used to fill the "time", "id", and "full text" fields of the object, similar in function to the three columns of the CSV table with the same name. Should the text of a quoted tweet contain another quoted tweet, the quote JSON object will store another quote object under the "quote" field. This assumes that each tweet can only quote a single other tweet, as per Twitter's Retweet guide [12].

If this is not the case, then no "quote" field will be included. A quote object without that field denotes the end of a chain of such objects. Note that the quote objects seen in the figures have expanded the JSON text with whitespace for clarity. The JSON objects have been stored on a single line to save space in the dataset. Data specifications, as outlined by the Elsevier Journal's Data in Brief format, is shown in Table1. Sample Tweets are shown in Table 2. Fig.1. illustrates the structure of a quote object in the JSON form with no sub quotes. Fig.2 shows the structure of a quote object in the JSON form with two recursive sub quotes.

Table 1. Data Specification

Subject	**Computer Science**
Specific subject area	Research on Twitter Sentiment and Bitcoin Price Correlation
Type of data	CSV Table
How data were acquired	Requested from Twitter's Search API
Data format	Raw, strings
Parameters for data collection	Tweets were collected between 8:46:33 AM, May 22, 2020 and 23:59:59 PM, June 22, 2020. Requested tweets contained at least one of the following keywords: bitcoin, Bitcoin, bitcoins, Bitcoins, BTC, XBT, satoshi, #bitcoin, #bitcoins, #XBT, #BTC, $XBT, $BTC
Description of data collection	Timestamps, id, and the text of truncated, retweeted, full, and/or quoted text from each tweet
Data source location	Online, requires internet access
Data accessibility	https://drive.google.com/drive/folders/1zB__72h1Q9N58diKzJ194B_Y-gaxycso?usp=sharing
Related research article	None

Table 2. Tweet rows for 5 sample tweets posted on June 22, 2020.

time	**id**	**text**	**original time**	**original id**	**original text**	**full text**	**quotes JSON**
2020-06-22 23:59:52	1275216...	Best ind...				Best ind...	{'time':20...
2020-06-22 23:59:51	1275216...	RT @IG...	2020-06-22 23:46:56	1275213...	#BlueLe...	#BlueLe...	
2020-06-22 23:59:50	1275216...	RT @Sm...	2020-06-22 6:55:33	1274959...	The fun...	The fun...	
2020-06-22 23:59:49	1275216...	@Alt...					
2020-06-22 23:59:47	1275216...	30 day...					

```
{
   'time': '2020-05-31 23:59:58',
   'id': '000000000000000003',
   'full text': 'Sample text for demonstrating the quote object's structure'
}
```

Figure 1. The structure of a quote object in JSON form with no sub quotes.

```
{
   'time': '2020-05-31',
   'id': '000000000000000002',
   'full text': 'Sample text for demonstrating the quote object's structure'
   'quote': {
              'time':'2020-05-31 23:59:58'
              'id': '000000000000000001'
              'full text': 'More sample text...'
              'quote': {
                         'time': '2020-05-31 23:59:58'
                         'id': '000000000000000000'
                         'full text': 'More sample text...'
                       }
}
```

Figure 2. The structure of a quote object in a JSON format with 2 recursive sub quotes.

IV. Experimental Design and Methods

The Tweet collection was done by a script made in Python. This script used the Tweepy library [13] to requests tweets through Twitter's Search API [8,9]. The script was designed to collect all tweets it could from 00:00:00 to 23:59:59 of the previous day and store them in a single CSV file. Making our own script allowed us to make any unforeseen modifications needed for maintaining access to Twitter. The collection is intended to be done for each day from the end of May 2020 to the beginning of August 2020. All tweets from 8:46:33, May 22 to 23:59:59, June 22 have been collected so far. The dataset will continue to be updated until August 7, 23:59:59, once tweets have been gathered for all remaining days.

V. Business Adoption

The collected BTC tweets are useful datasets for many business applications to predict the BTC market movement directions. Financial analytics [14]-[16], anomaly detection [17]-[19], and machine and deep learning [20]-[22], and association mining [23],[24]..

VI. Conclusion and Future Directions

In this paper, we have provided the NLP community with a new twitter corpus collected in the era of COVID19. This corpus can be used in text mining, price prediction, financial analytics, and trend prediction. Future Directions include extending the corpus size to a more extended period to collect tweets in the entire COVID-19 period.

ACKNOWLEDGMENTS

This research has been supported by the Ryerson University, Faculty of Engineering and Architectural Science Undergraduate Research Opportunity Fund

REFERENCES

[1]. E. Stenqvist and J. Lönnö, 'Predicting Bitcoin price fluctuation with Twitter sentiment analysis', Dissertation,. Which journal it was published in ?, School of Computer Science and Communication, 2017.

[2]. O. Kraaijeveld and J. D. Smedt, "The predictive power of public Twitter sentiment for forecasting cryptocurrency prices," Journal of International Financial Markets, Institutions and Money, p. 101188, Mar. 2020.

[3]. T. R. Li, A. S. Chamrajnagar, X. R. Fong, N. R. Rizik, and F. Fu, "Sentiment-Based Prediction of Alternative Cryptocurrency Price Fluctuations Using Gradient Boosting Tree Model," Frontiers in Physics, vol. 7, Oct. 2019.

[4]. S. Mohapatra, N. Ahmed, and P. Alencar, "KryptoOracle: A Real-Time Cryptocurrency Price Prediction Platform Using Twitter Sentiments," Feb. 2020. Which journal it was published in ?

[5]. C. Kaplan, C. Aslan, and A. Bulbul, "Cryptocurrency Word-of-Mouth Analysis via Twitter," ResearchGate, 2018. [Online]. Available:

[6]. A. Jain, S. Tripathi, H. D. Dwivedi and P. Saxena, "Forecasting Price of Cryptocurrencies Using Tweets Sentiment Analysis," 2018 Eleventh International Conference on Contemporary Computing (IC3), Noida, 2018, pp. 1-7, doi: 10.1109/IC3.2018.8530659. Conference Name:ACM Woodstock conference

[7]. https://developer.twitter.com/en/docs/basics/things-every-developer-should-know

[8]. https://developer.twitter.com/en/docs/tweets/search/overview

[9]. https://developer.twitter.com/en/docs/tweets/search/overview/standard

[10]. https://developer.twitter.com/en/developer-terms/agreement-and-policy

[11]. https://developer.twitter.com/en/docs/basics/counting-characters

[12]. https://help.twitter.com/en/using-twitter/how-to-retweet

[13]. https://github.com/tweepy/tweepy

[14]. Xue Tan and Rasha Kashef. 2019. Predicting the closing price of cryptocurrencies: a comparative study. In Proceedings of the Second International Conference on Data Science, E-Learning and Information Systems (DATA '19). Association for Computing Machinery, New York, NY, USA, Article 37, 1–5. DOI:https://doi.org/10.1145/3368691.3368728.

[15]. Ibrahim, A.; Kashef, R.; Li, M.; Valencia, E.; Huang, E. Bitcoin Network Mechanics: Forecasting the BTC Closing Price Using Vector Auto-Regression Models Based on Endogenous and Exogenous Feature Variables. J. Risk Financial Manag. 2020, 13, 189.

[16]. Tobin T., Kashef R. (2020) Efficient Prediction of Gold Prices Using Hybrid Deep Learning. In: Campilho A., Karray F., Wang Z. (eds) Image Analysis and Recognition. ICIAR 2020. Lecture Notes in Computer Science, vol 12132. Springer, Cham. https://doi.org/10.1007/978-3-030-50516-5_11.

[17]. Kashef R., Gencarelli M., Ibrahim A. (2020) Classification of Outlier's Detection Methods Based on Quantitative or Semantic Learning. In: Fadlullah Z., Khan Pathan AS. (eds) Combating Security Challenges in the Age of Big Data. Advanced Sciences and Technologies for Security Applications. Springer, Cham. https://doi.org/10.1007/978-3-030-35642-2_3.

[18]. Kashef, Rasha, and Mohamed S. Kamel. "Towards better outliers detection for gene expression datasets." 2008 International Conference on Biocomputation, Bioinformatics, and Biomedical Technologies. IEEE, 2008.

[19]. Rasha F. Kashef ; Proceedings of the KDD 2017: Workshop on Anomaly Detection in Finance, PMLR 71:43-55, 2018.

[20]. Kashef R. (2020) Adopting Big Data Analysis in the Agricultural Sector: Financial and Societal Impacts. In: Pattnaik P., Kumar R., Pal S. (eds) Internet of Things and Analytics for Agriculture, Volume 2. Studies in Big Data, vol 67. Springer, Singapore. https://doi.org/10.1007/978-981-15-0663-5_7

[21]. R. Kashef and M. S. Kamel, "Hard-fuzzy clustering: A cooperative approach," 2007 IEEE International Conference on Systems, Man and Cybernetics, Montreal, Que., 2007, pp. 425-430, doi: 10.1109/ICSMC.2007.4413889.

[22]. A. Ibrahim, D. Rayside and R. Kashef, "Cooperative based software clustering on dependency graphs," 2014 IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE), Toronto, ON, 2014, pp. 1-6, doi: 10.1109/CCECE.2014.6900911.

[23]. El-Sonbaty, Yasser, and Rasha Kashef. "New Fast Algorithm for Incremental Mining of Association Rules." ICEIS (1). 2004.

[24]. Sonbaty, Yassar El, and Rasha F. Kashef. "NBP: Negative Border with Partitioning Algorithm for Incremental Mining of Association Rules." (2004).

Static Analysis Framework Based on Multi - Agent System

Haissam Badih, Yasamin Alagrash
Department of Computer Science And Engineering
Oakland University
Rochester, Michigan 48309, USA.
{hbadih, yhalagrash}@oakland.edu

***Abstract*—The progress of the methodology of static analysis has recently been rapidly consistent. Advancement is gradually attracting debugging and detailed programs by applying the intelligent approach to static analytical techniques. This paper introduces a novel framework for static analysis techniques based on a multi-agent system. The work investigates an innovative approach for multiple static analysis methods in an intelligent way under a distributed environment. To verify the proposed framework, we implement the traditional method of static analysis based on multi-agent services and determines a head track for autonomous and efficient static analysis techniques. Our approach has used agent characteristics in a multi-agent context. All agent services communicate with one another, adjustments were made to the environment of the agent, various restrictions were applied, and the systems' effects were calculated. In order to demonstrate the respective method's strengths and weaknesses, each system has been analyzed using several performance metrics. This paper develops multiple static analysis is applied to enhanced system reliability and scalability. Theses provide server-side static analyzes. Our solution shows high performance within a reasonable time duration.**

***Index Terms*—Multi-agent system, distributed static analysis, intelligent static analysis, Intelligent system**

I. Introduction

Web Services (WS) applications have unique characteristics that need to be, ponder upon to achieve successful and optimal integration of the applications. Such characteristics vary from run-time behavior to the collaboration among the developers. Service-oriented architecture (SOA) has fortunately provided an open architecture for the integration of WS applications of this nature. This architecture is loosely-coupled and standard-based. Now, as the architecture provided by SOA is an open platform, services can be published at any time, which means there will be so many services serving the same purpose. The problem arises when an application switches to another service during system evolution, and the system should be consequently reconfigured at run-time [1]. In recent years, several attempts try to build scalable program analysis tools, more scalability, and self-depend analysis entity come from development in Setup Analysis Tool (SAT) solver, data analysis techniques, and data log [4]. Traditional program analysis is designed to work under a single node or machine within the limits of data storage and input data size [3]. This work introduces to novel static analysis approach deals with big data over a distributed environment. This paper discusses the static analysis of the multi-agent system to follow all scalable and distributed static analysis requirements.

Generally, an agent is an autonomous and self-dependent entity that can interact on its own and perform tasks. Similarly, a group of agents can be combined to better coordinate and orchestration the best results [2].

II. Static analysis based on Multi- Agent System

Static analysis is an automated technique that explores program properties; it has been helped in code debugging technique, program understanding, and optimization, which involves code inspection without executing it [13]. Agents Smarter entities within a society can change the interaction between other entities in the same society. The Agent will be the brain of a society, which controls the interaction between different agents. In such a way, if the core agent sends a request to an agent in the society, the core continues or discontinues based on the other Agent's response.

This way, Agent will manage the traffic of the communication between the agents, better than a Multi-Agent system. Also, the core agent can easily build a benchmark for any transaction in society. Multi-agent society is a society of software modules with socialization capabilities with two types of behaviors: first is the individual behavior that launches from knowledge base built upon the creation of the entity and reforming occurred due to the interaction with the environment. Second is the society behavior, which is the dominant behavior [12].

In simulating human beings, people tend to grant decisions come from known smarter members high credibility (i.e., weight) while others are ranked lower in term of weights, in this paper. We will examine developing knowledge in the presence of smarter agents; those agents are no different than others, but they have a more extensive knowledge base and more experience in reaching conclusions faster. Besides, it is very costly as well as time-consuming to do static analysis for the distributed systems.

However, as MAST is composed of groups of intelligent agents, they carry out the static analysis faster and with more efficiency. The agents monitor the behavior of the distributed system and analyze where testing is necessary. All agents communicate with each other and autonomously discover the defected areas in the distributed system. The agents can

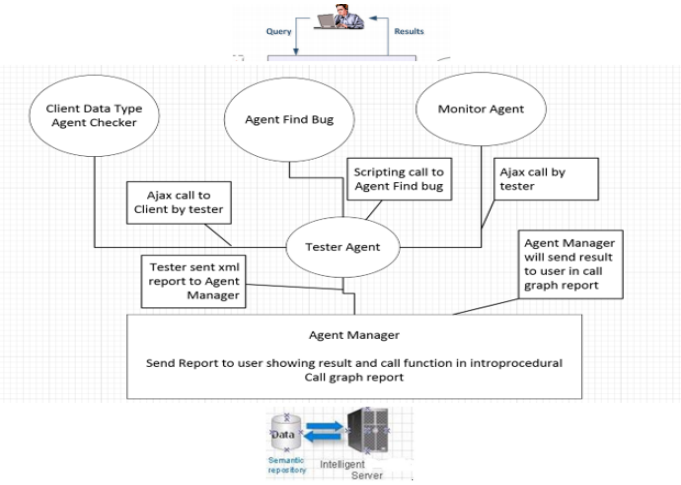


Fig. 1. Architecture of our approach to apply multi-agent system in static analysis.

carry out all static analysis techniques, for instance, data flow analysis, variable type [2]. All of these techniques make the distributed system least vulnerable to errors or defects.

Novelty: - this work based on Multiple static analysis based on multi-agent system proves to increase the efficiency and can be re-used. Using multi-agents, miscellaneous types of static analysis can be done at the same time. Moreover, intelligent agents ensure that the static analysis completes within a minimum timespan regardless of the program's size or complexity. These agents extend and further advance static analysis techniques for WS applications. Along with the dynamic reconfiguration and recomposed agents, various levels of static analysis, such as front-backend, can be achieved. This paper introduces a novel multi-agent static analysis framework to Increase the efficiency of static analysis techniques.

Organization: The remainder of this paper is organized as follows.In Section III describes our approach to apply multi-agent system. In Section IV we discuss the testing and validation of our approach. Section V we discuss related works on using agent and multi-agent in static analysis Section VI summarizes our findings and concludes the paper.

III. Approach and Methods

The distributed system has a simple structure consisting of a server, middleware, and multiple clients. A distributed application started between them, which contains a server data repository. The middleware provides the connection bus between the server and the clients. Figure 1 shown the proposed system structure. In this work, we intend to build JAVA agents that, autonomously, are interacting web environment; those agents are trained with Mentioned JAVA agents. They will be implemented as a web service; the web application is built to host instances of JAVA agents.

The client interface can be validated from an agent validator to meet the required input. If an agent requests a call to the agent manager to check whether their request has been existing or not, they will respond to the manager. All the perceived objects and the perceiving schemes will be saved in the semantic repository to reduce the heavy load of processing and network traffic. The distributed systems can use static analysis efficiently with the help of multiple intelligent agents that are autonomous and monitor the behavior first and act accordingly collaboratively.

A. User request process motivation example

Since we are using different agents to do the testing, we need to give the user a result—this result base on request and response. The user will make a request, and then the program will then provide the result back as a response. This response will be in the call graph. The call graph is the control flow graph, which characterizes the calling relationship between function and program. Each node represents a function or method, and each edge will represent that method M1 call G1. Therefore, the cycle of a graph indicates what function or method is calling which method.

This very important to figure out an analysis like you want to make sure the right type inference you want to introduce the generic type in the right point. Call graphs are a primary program analysis result that can be used for human understanding of programs, or as a basis for further analyses, such as an analysis that tracks the flow of values between procedures. A straightforward application of call graphs is finding procedures that are never called. Call graphs can be dynamic or static analysis.

A dynamic call graph is a record of the program's execution, for example, as output by a profiler. Thus, a dynamic call graph can be exact, but only describes one run of the program. A static call graph is a call graph intended to represent every possible run of the program. The exact static call graph is an undecidable problem, so static call graph algorithms are generally over approximations. Every call relationship that occurs is represented in the graph, and possibly also some call relationships that would never happen in actual runs of the program.

Additional details about how agents communicate. We know today there are lots of automation testing tools, usually, the automation testing tools are built on the framework, like Selenium and others like this framework. But this will allow us to have an automation testing to see where the bugs are happening, causing the application's performance issues, etc. This all good and excellent. But the idea is to create specific agents to test the app to see the static analysis of the code. It doesn't matter what environment we are using. Our purpose here is to find the bugs and warning of the application from static analysis of the code by using our framework.

The tester agent will make Ajax call to the client agent, and then the client agent will send the result to the tester. After the client agent did the tester will continue base on the outcome. This result will be sent to the Manager agent, and if negative, the agent manager will send a report back to use in the call graph to identify the errors. If the result positive from the client, the tester will continue to process the find bugs. If the

find bug agent has a negative request, it will send the report Manager agent, and this agent will send the result to the user after receiving the XML report from the tester, and this XML will be transferring to call graph as a friendly user.

Suppose the agent finds bugs have a positive result. In that case, it will continue to do the test automation, and the Monitor agent will send a response back to the tester. The tester will send an XML report to the Manager agent. The manager agent will generate the result in the call graph showing every function call result base intraprocedural. That will be reported back to the user in the call graph report. After all agents complete the test, it will return all automation results in a call graph report showing the intrprocedural flow to the user, including the class and function, including all exceptions from the monitor agent.

The call graph is to generate a report to the user to show the result in friendly users, there are frameworks out there and one of them called soot can be used to generate reports. That can visualize the result in the flowchart view.

Since we are using different agents to do testing, we need to give the user a result—this result is based on request and response. The user will make a request, and then the program will provide the result back as a response. This response will be in the call graph.

B. Challenge Technique:

JADE is (Java Agent Development Framework) is a software Framework fully implemented in the Java language. It simplifies the implementation of multi-agent systems through a middleware that complies with the Foundation for Intelligent Physical Agents (FIPA) specifications and through a set of graphical tools that support the debugging and deployment phases. A JADE-based system can be distributed across machines (which not even need to share the same OS), and the configuration can be controlled via a remote GUI.

The configuration can be even changed at run-time by moving agents from one machine to another, as and when required. JADE is completely implemented in Java language, and the minimum system requirement is version 5 of JAVA (the run time environment or the JDK). In addition to the agent abstraction, JADE provides a simple but effective task execution and composition model, peer to peer communications based on the paradigm passing an asynchronous message, a framework for publishing subscription discovery, and many other advanced features to promote the development of a distributed network. The startup agent, as shown in Figure **??**.

It will trace the communication between agents between request and response, and we have two agents and communicate with each other since this only for communication with agents and has to be with the same environment. And if we see, two agents communicate with each other. The main container will show the agent and in it can call the sniffer to do a sniff for each agent, and it will show the request and respond between the agent, therefore, since Jade is to show communication between the agent and show the request and respond 4.

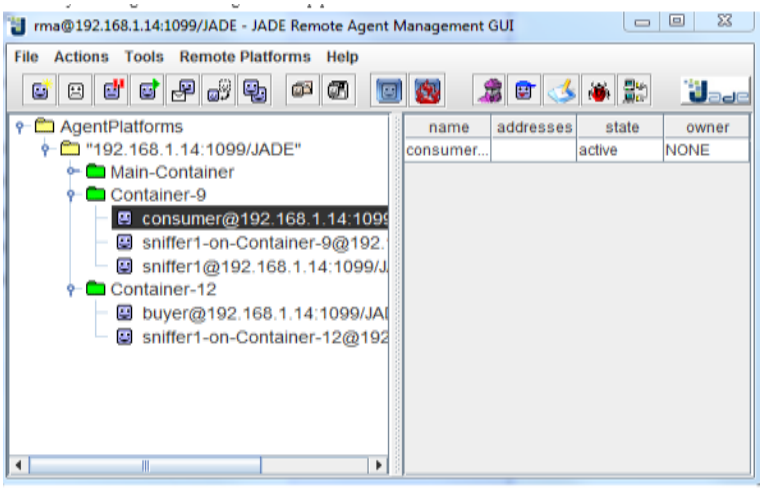


Fig. 2. Agent Environment

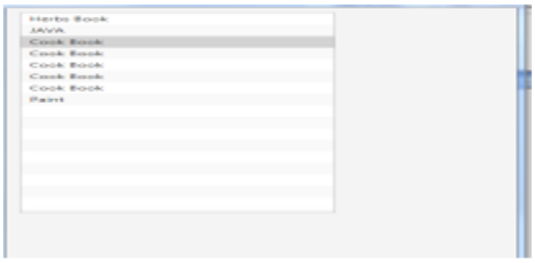


Fig. 3. Agent Creation

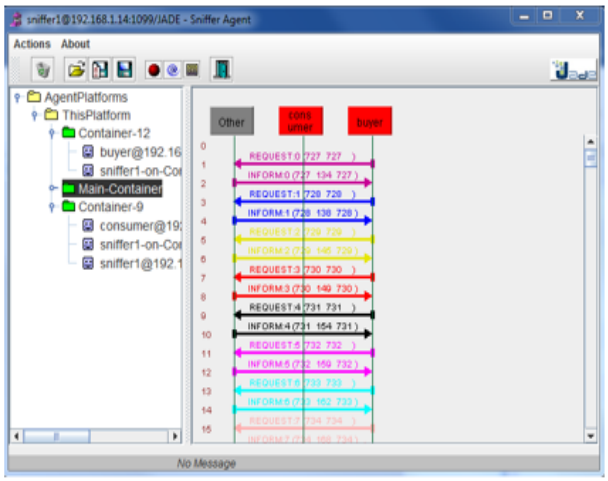


Fig. 4. Agent Creation

We develop the framework for static analysis for a different environment. Since we are using different agents to do testing, we need to give the user a result—this result base on request and response. The user will make a request, and then the program will provide the result back as a response. This response will be in the call graph. The call graph is known as the control flow graph, which characterizes the calling relationship between function and program.

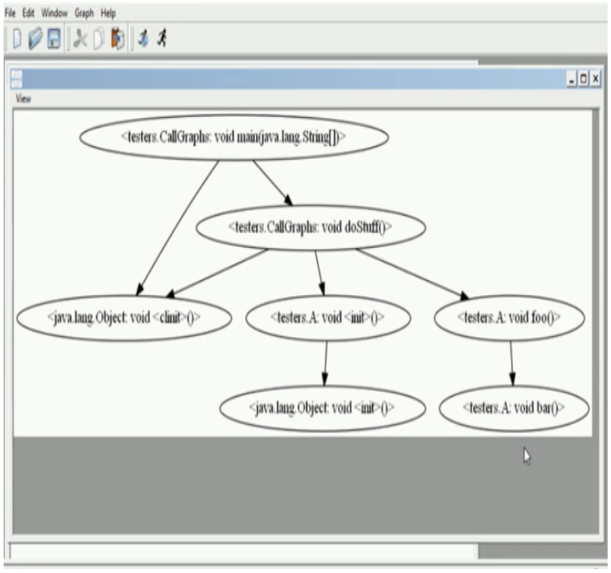


Fig. 5. Agent Graph

Each node represents a function or method, and each edge will represent that method M1 call G1. Therefore, the cycle of a graph indicates what function or method is calling which method. This very important to figure out an analysis like you want to make sure the right type inference you want to introduce the generic type in the right point. So, Call graphs are a basic program analysis result that can be used for human understanding of programs, or as a basis for further analyses, such as an analysis that tracks the flow of values between procedures. A straightforward application of call graphs is finding procedures that are never called.

Call graphs can be dynamic or static analysis. A dynamic call graph is a record of execution of the program, for example, as output by a profiler. Thus, a dynamic call graph can be exact, but only describes one run of the program. A static call graph is a call graph intended to represent every possible run of the program. The exact static call graph is an undecidable problem, static call graph algorithms are generally over approximations. That is, every call relationship that occurs is represented in the graph, and possibly also some call relationships that would never occur in actual runs of the program.

Therefore, our agent will send the response to the user in such a call graph result. That will indicate all nodes and edges that the call did go through. This way will identify the analysis of the program or the functionality of the request. Using dataflow analysis and Introprocedural flow analysis 5.

Static Analysis Services: - For our Static Analysis Services, we need flow analysis for each agent and the agents are:

- Client Parser Agent
- Tester Agent
- Find Bug Agent
- Monitor Agent
- Manager Agent

Algorithm 1: Perception algorithm for agent's static analysis.

```
Input: Sensor capture data
Result: Result action output.
while Not Null do
    Sensory input;
    Activate modeling Layer;
    Activate planing layer;
    Activate reactive layer;
end
```

This algorithm is a generic algorithm that is working in every agent. The perception subset system will be working differently from every agent. Focus on the role of the agent in a global context and the function of the agent. More specifically, we explain each step within each agent as a functional step that triggers automatically to control the system architecture.

How the sensor is working with webservice agent? Basically, sensor is going to sense the environment. To identify the structure of our sensor function we follow the following structure: 1 Agent 2 option 3 service name

Form the data; we can initialize the sensor option. From the option, we can identify the configuration needed, like service name, port number, and data passing through. The initial sensor initializes the sensor (without tracing) to begin collecting and reporting information. For example, if we have data for login, we can identify which service required to go.

All agent's services interact with each other to have the interaction we need to use a flow analysis to be able to have a good static analysis. In the client parser service and find bug service will be using data type analysis. The tester service will be using interprocedural data flow analysis from monitor service to trace the tester service. Manger service will be the repository of the result of the trace. Table I shown functional of agents.

TABLE I
STATIC ANALYSIS AGENTS

Agent Name	Goal	Function
Agent Client Parser	Validate Data	Capture data and Analyze it
Agent tester	Validate functionality	Result test
Agent find Bug	Validate Object	Log file
Agent Monitor	Since the repository call	Identify services
Agent manager	Capture all result	Save to database global

Algorithm 2: Algorithm of agent's static analysis.

```
Input: user request submitted.
Result: Program and system analysis.
while Not Null do
    tester agent will receive a request;
    the tester identify which agent needed;
    call specific agent;
    specific agent will request call to manager agent for an existing scenario;
    if scenario exist then
        respond to user;
    else
        specific agent will send report to tester agent;
    end
end
```

IV. Testing And Validation

We need flow analysis for each agent, and the agents are 1.Client Parser Agent 2 for our Static Analysis service. Agent Tester 3. Find Agent 4 for Bug. Agent 5 track. Manager Agent, The services of all agents, interact to ensure interaction between us to use the flow analysis to have a functional static analysis. Data type analysis is used in the client parser service, and the bug finds service. The trace will be used in the test service by interprocedural tracking data flow analysis. The software Manger will be the trace results server. Functional agents display in I.

A. Client Parser Agent

This service will be for the client site to check the correct type of the input and to do that we will be using the Data Type analysis. What is data type analysis and how it works?

B. Data Type Analysis

Type is having been around as long as programming language themselves. Their two motivations for type one type will be for what type of data representation for a particular piece of information and the other type to qualify the logical properties of the values we need to operate on. Type is, in general, an idea of a contract of what variable can contain.

For example, if we have v: T it means v it is an expression of type tie T, and this is a type for v, and it will be in this type and can do in this type which can not do other types. Figure 6 shows the expression and type for the analysis. When we have type in the code, it acts constructively, and it won't break each other. And this leads to a notion of a strong typing it leads there is no memory error like if we read a type integer from somewhere, that means there is an integer from somewhere. And by the way, not all language has type safety like Assembly and weakly type safety like C and C++ and strong type safety like Java and Python.

Strong typing comes in two flavors static type and dynamic typing. Figure 6 shown a motivation example of static analysis using Agent parser Because x is declaring as a type of integer. In the program assigned type of string will see the compilation error. In our agent services, the one they use the data type analysis will be looking for identifying errors during compilation time and identifying the programmer where the errors are located. Why the errors occur like, for example, if we are using the data type for client parse in the javascript, we will identify the wrong type inserted in Graphical User Interface (GUI). Therefore, data won't be transfer to the middle tier, and this can be run by a script to identify the errors.

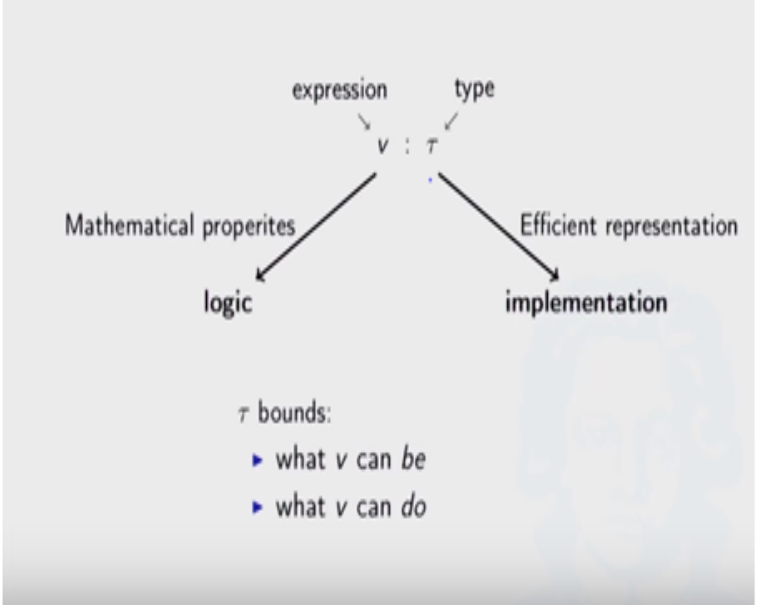


Fig. 6. Data Type Analysis

For the server code, we will inject a third party tool to identify java code error if a variable declare or an object declare and access object method and if the object instantiated not null. figure 7 shown front-end variable type analysis. All this will be in log files after the script been executed. Also, we will be treating dynamic analysis; for example, if we have errors happen during runtime, it will be identified before the run by using third party tools like finding a bug or other tools. In Summary: Types are contracts over typed values: Guaranty logical properties: Inform implementation / in-memory representation Type safety describes the rigor the language employs, from un-typed to strongly typed. Strongly typed: detect all type errors before they happen. Statically typed: detect type errors at compile time. Dynamically typed: detect type errors at runtime by using the third party can identify the error and be logged as shown in Figure 8.

C. Tester Agent:

The main reason for this Agent service is to access the client parser and find bug service by using these services to check the static type and dynamic type before executing the code. After code passed by the client parse and find a bug, the tester will have to execute different codes by using web services; therefore, the tester will be using Interprocedural Data Flow. As we mention in the system overview, the tester will be built from many functions.

The first function will be by calling the client parser to check the data type, and the second function will be by calling the found bug to identify the code is right by compilation, and no instant object will be null in the run time, and that will be by a warning. The other functions will be accessing different web services functions, and these services will trace by Interprocedural data flow.

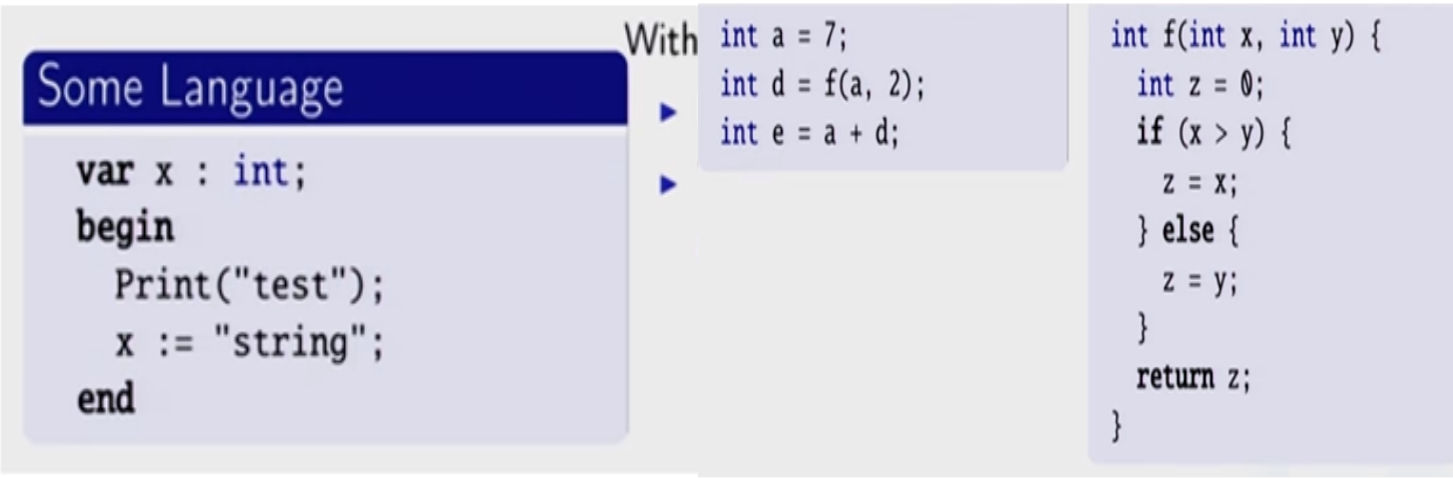


Fig. 7. Motivation.

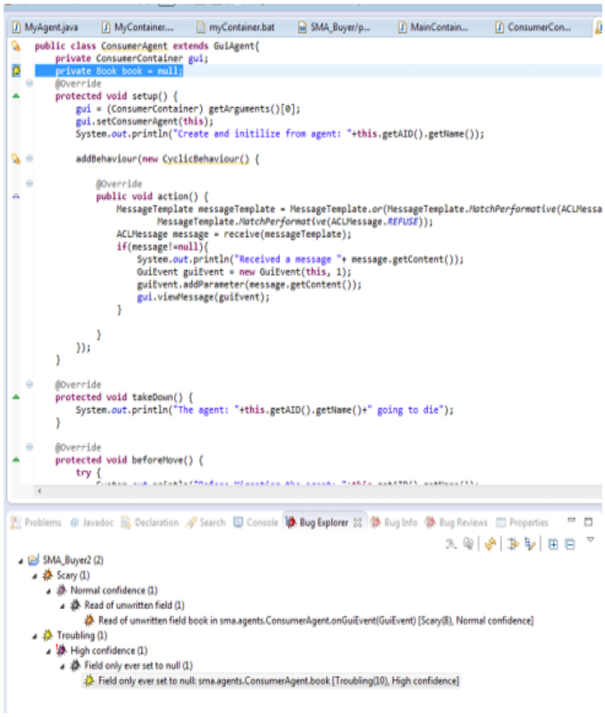


Fig. 8. Static Analysis.

We will be looking for Interprocedural Data Flow considering the following example we have a program in the left and functional procedural on the right. We have an integer a and integer d call a subroutine f(intx, int y) then we will add the result with a and assigned to e, we notice our analysis does not know what have to do with f the only thing, it will assign something to it. This will be imprecise and unfortunate. If we have to implement a definition like what we have, we need to be a little smarter, as shown in Figure 9.

We need to take the program and break it into a node with edges, and these will be the control edges. We will have two graphs, one on the right-hand side will be the program, and

Fig. 9. Static Analysis.

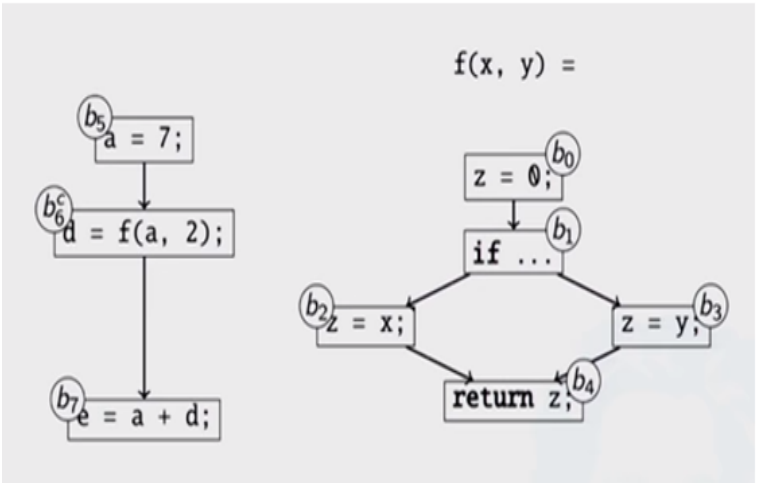


Fig. 10. Static Analysis.

the left-hand side will be the subroutine f; the f is a simple subroutine with simple conditions. And in the left-hand side is linear program simple, also as shown in Figure 10.

We add edges to specify the call function are happening the first we will show is from the call side to the first node of the f function. So, the f call will be passing the values to the f function x will be for a and the x could be seven and y is two as we see in the figure above. Since we have an edge from the call to f in the first node, then we need another node for the return, as we are showing in below graph 11. And the separate node here is annotate to whatever result we will get and to catch the analysis result. We add the separate node here that allows us to join all the things the f is not changing.

The transfer function from the call function bc6 to br6 is going to tell us the assignment of a, and we can join with assignment coming from the return that tells d might be 2 or 7. So to join the result in br6 the out from b6 and the out from b4 will give the out of b6r and then the d will be coming from the return node and the result will be add in b7 and the

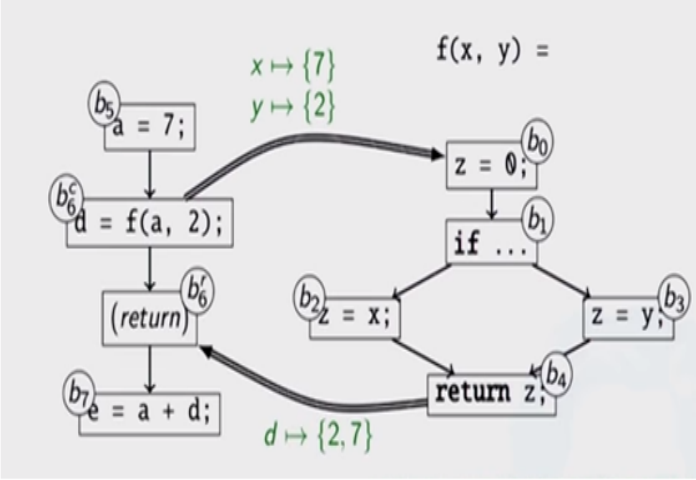


Fig. 11. Static Analysis.

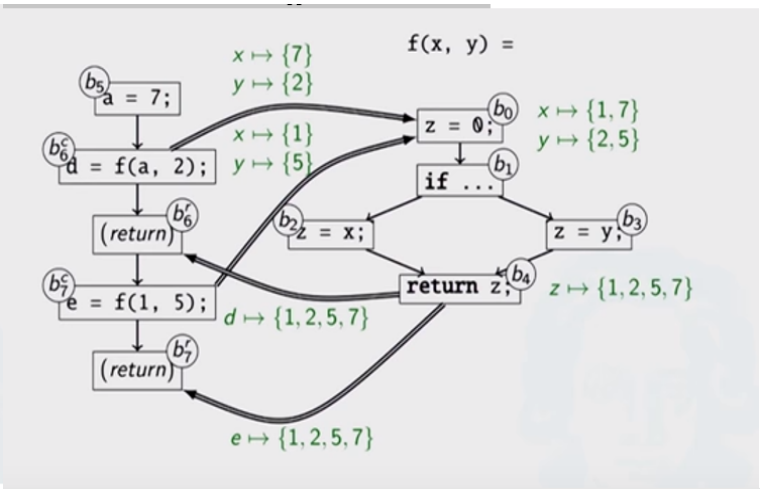


Fig. 13. Static Analysis.

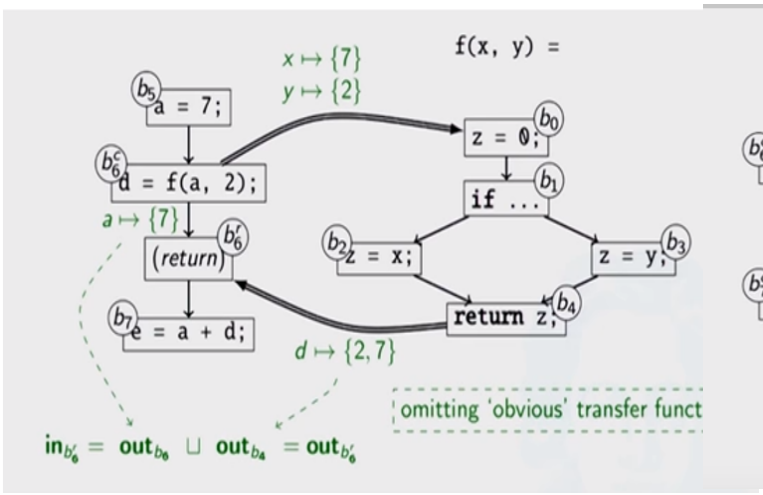


Fig. 12. Static Analysis.

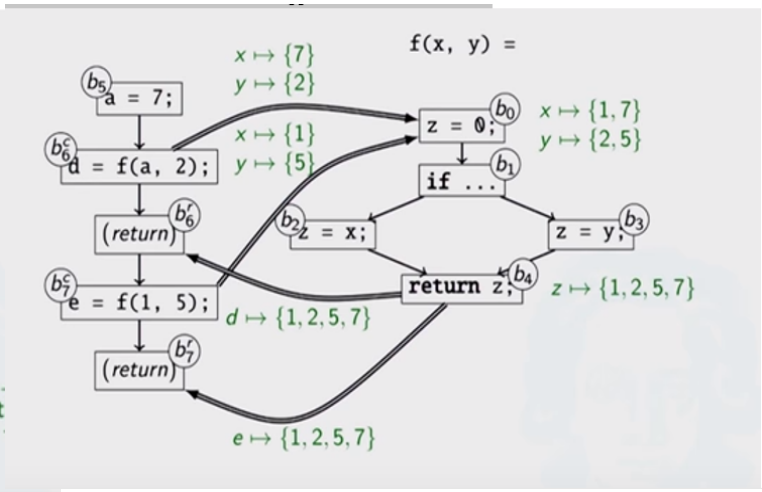


Fig. 14. Static Analysis.

out b7 can be [14], [5] because that's what the single possible assignment is going to tell us. So that nice and useful result so far 12. But let's make things a little interesting and what we mean by that is we are going to introduce the second function.

The image below will show how to call the second function in Figure 10. And we are going to make the same approach we did for one function call. We will introduce the return for both function and the call edges for both function calls into the beginning of f, now through the calls is going to pass the concrete parameters for the x and y for the first call will be 7,2 and the second will be 1,5 as shown in Figure 11.

What self-respecting data flow has to do when it has multiple inputs is to join the input. We will apply the data flow, and what we can get the return values might be 1,2,5,7, and the analysis d return it might be 1,2,5,7, and e might be 1,2,5,7. This analysis is perfectly sound and covering all sorts of results that can happen, but kind of imprecise thy are not context-sensitive, and this is a limitation of the approach we looked so far as showing in Figure 14. So, what we have gone though:

- Split call sites bx into call (bcx) and the return (brx) nodes.
- Intra-procedural edge bcx -¿ brx carries environment/store
- Inter-procedural edge (-¿)
- Caller -¿ subroutine, substitutes parameters (for pass-by-value)
- Caller ¡- return, substitutes result (for pass-by-result)
- Otherwise as intraprocedural dataflow edge
- Naive approach(Consider so far):
  - All call sites connected to subroutine entry
  - Subroutine return connected to all return sites
  - Context-Insensitive

We want to have Context-Sensitive, and we need to go back to the interoprocedural we did before and try to come up with a good Context-sensitive. So we need to go back to our notion of paths that we introduce with MLP algorithm, and we extend this notion to call paths, and the reason we do that is that we

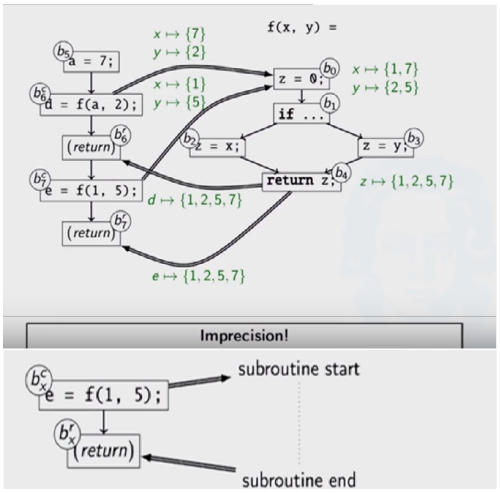


Fig. 15. Static Analysis.

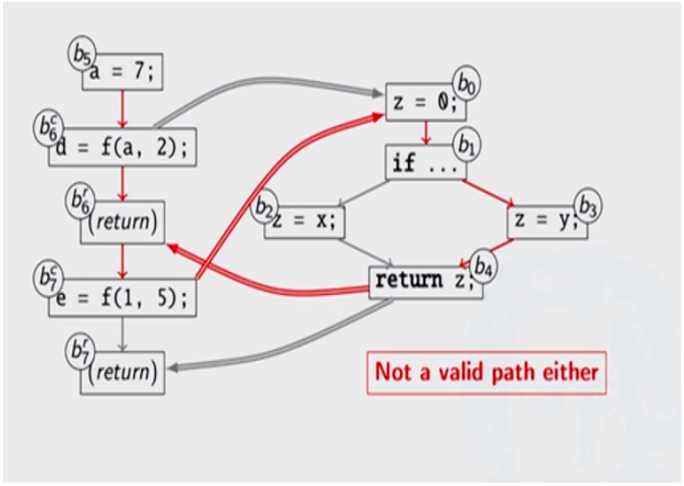


Fig. 16. Static Analysis.

want to avoid analyzing invalid paths. The invalid paths are after all what is missing things up here invalid paths might be going from call node from node b5 to bc6 then going to the entire function then returning to node b7r that is the exact case when you take some information that is flowing from one call and returning it to different unrelated call so this is something that we want to disallow. Similarly, we want to disallow the opposite that if we are at point b7 in the program and we are calling the subroutine that we return the results for that b7 node to b6 node those paths. As shown in the figure as shown in figure 16

D. Context Sensitive

We want to exclude but we want to include all the valid ones so all the ones can happen during execution and for those we are going to look at paths For those we are looking at paths going from [b5 ,bc6, b0, b1, b3, b4, br6]

V. Related Works

Several generic approaches work into a multi-agent system to test a distributed system. Papers [7], [4] used test agents to make it easy and bring out flexibility, ease, and extensibility in testing web applications using any framework. Generally, there are four-level dataflow testing approaches along with four-level test agents in agent-based testing. Some basic agent-based testing approaches can be implemented to all kinds of web applications, but specific test agents are classified for specific web documents. Our work extended Inter-procedural data flow Analysis through call graph techniques with multiple data type analysis methods.

Well Qi et al. [1] propose that it's necessary to integrate various types of testing approaches to bring out the systematic testing of the web applications. This work focuses on testing strategy with no approach for static analysis. Yamany et al. [2] propose a multi-agent framework for testing a three-tier distributed system. The framework promises to bring out the ease and reliability in the testing phase.

The framework inculcates intelligent agents that track user's actions or activities, generating new test cases accordingly. Also, the agents help generate an expected output of the execution of certain test cases, consequently helpful invalidating the test cases. Yamany et al. propose a scenario where a three-tier-web-application can be tested in three different cases.

The scenario validates the framework and helps in extracting the errors at each layer of the web application. The scenario also extracts the integration-level errors (which occur while integrating three layers of the web application). The use of intelligent agents for testing web applications is a step towards automated testing, which reduces time and cost. But, this work is doing testing on only web services.

Garbervetsky et al. [16] proposed a distributed static analysis framework that can test the application according to the size of the input. The framework makes use of a distributed analysis framework, which in turn demonstrated how well it could do testing on a large volume of data without getting short memory. The framework demonstrated an impressive use of static analysis on the distributed system, which further uses call graph analysis. However, the system does not deal with updates in code and throughput in increment analysis. Our framework fixed this issue by using Inter-Procedural Dataflow Analysis techniques based on a multi-agent system.

Not many papers in static analysis adopt a multi-agent system in the analysis process. Most articles deal with test agents in a distributed system. However, few approaches applied distributed static analysis within dataflow Analysis [16]. This paper develops distributed static analysis is applied to enhanced system scalability.

VI. Conclusion

To address the challenges of collaborative services-oriented static analysis, this paper displays a framework for static analysis with multi-agent system. Multiple static analysis tools are using to develop and extend these techniques. Our framework is re-composed and re-configuration automatically

based on using smart agent services. Time and cost criteria are considering in our framework. The analysis will complete in a reasonable time. No matter how big a program is given.

Our research introduces a robust approach to software engineering area, which can conduct static analysis with less complexity within a short time. Our work is the first work compared to other previous results, to develop a multi-agent system for call graph analysis during a runtime error with program analysis. In future work, we plan to design and implement a multi-agent system with static analysis as Software as a Service (SaaS) in the cloud computing environment; this implementation provides a platform to access the system remotely by clients' Agents everywhere'.

REFERENCES

[1] Chen, Y.S., 2010, April. The Generating Algorithm of Protocol Consistency Test Sequences Between Combined Control Flow and Data Flow. In 2010 International Conference on Machine Vision and Human-machine Interface (pp. 480-482). IEEE.

[2] Barenji, A.V., Barenji, R.V. and Hashemipour, M., 2016. Flexible testing platform for employment of RFID-enabled multi-agent system on flexible assembly line. Advances in Engineering Software, 91, pp.1-11.

[3] Orso, A., Harrold, M.J. and Vigna, G., 2001. MASSA: Mobile agents security through static/dynamic analysis. In in Proceedings of the ICSE Workshop on Software Engineering and Mobility (ICSE2001.

[4] Huo, Q., Zhu, H. and Greenwood, S., 2003, November. A multi-agent software engineering environment for testing Web-based applications. In Proceedings 27th Annual International Computer Software and Applications Conference. COMPAC 2003 (pp. 210-215). IEEE.

[5] Reps, T., Horwitz, S. and Sagiv, M., 1995, January. Precise interprocedural dataflow analysis via graph reachability. In Proceedings of the 22nd ACM SIGPLAN-SIGACT symposium on Principles of programming languages (pp. 49-61).

[6] El Yamany, H.F., Capretz, M.A. and Capretz, L.F., 2006, September. A multi-agent framework for testing distributed systems. In 30th Annual International Computer Software and Applications Conference (COMPSAC'06) (Vol. 2, pp. 151-156). IEEE.

[7] Kung, D., 2004, September. An agent-based framework for testing Web applications. In Proceedings of the 28th Annual International Computer Software and Applications Conference, 2004. COMPSAC 2004. (Vol. 2, pp. 174-177). IEEE.

[8] Beyer, D., Gulwani, S. and Schmidt, D.A., 2018. Combining model checking and data-flow analysis. In Handbook of Model Checking (pp. 493-540). Springer, Cham.

[9] Wehrli, S., 2017. Static program analysis of data usage properties (Doctoral dissertation, Master's thesis, ETH Zurich, Zurich, Switzerland).

[10] Pandey, A.K. and Pandey, R., 2015, March. Role of multi agent system methodology in system design. In 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 438-441). IEEE.

[11] Buguillo, J., 2018. Self-organizing Coalitions for Managing Complexity. Springer International Publishing, 29, pp.89-100.

[12] Khorsheed, A., 2014. Automatic Image Annotation using Image Clustering in Multi–Agent Society.

[13] R. Christoph. 2016 Dataflow analysis.

[14] Cox, D.R., 2018. Analysis of binary data. Routledge.

[15] Pandey, A. K., And Pandey, R. (2015, March). Role of multi agent system methodology in system design. In Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference on (pp. 438-441). IEEE.

[16] Garbervetsky, D., Zoppi, E., And Livshits, B. (2017, August). Toward full elasticity in distributed static analysis: the case of callgraph analysis. In Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering (pp. 442-453). ACM.

Audiovisual Interactive Companionship; The Next Breakthrough in Computers

Shahriar Khan
Dept of Electrical and Electronic Engineering
Independent University, Bangladesh
Dhaka, Bangladesh
skhan@iub.edu.bd

***Abstract:*--New developments in voice and facial video generation suggest we are at the verge of a new breakthrough in audio and visual interactive companionship. Today's computer-interactive voices (*Alexa* and *Google Assistant*) are largely repetitive, robot-like, and all-knowing, sometimes making them tiring and monotonous. The challenge is to better simulate the human voice and personality, with a face generating the voice. Covid and sprawling cities are making people increasingly isolated. A game or program on TV can be better enjoyed when accompanied by a computer companion. Other than companionship for entertainment, there can be companionship for therapy, such as at a hospital or at an old-age home. Personalities of famous people such as scientists, statesmen, actors, and sportsmen can be recreated. The computer voice can be a training ground for social etiquette for children and adults. A simulated baby voice can be used for training would-be parents. A simulated patient's voice can be used for training doctors and nurses. With videoconferencing being the new norm in these times of Covid, the simulated voice and video will appear more real than before. The bonding of humans with the computer voice raises ethical questions about whether this could become addictive, and whether details of the interaction can be used by the company and the government. Ominously, can the simulated voice of a child become the face of the government, using the data to profile or even arrest users? Could the government use the voice to inspire users to be hard working, law abiding, and tax-paying citizens? It is expected that market forces will prevail, and companionship with the computer voice and video will prevail. We must find ways to regulate computer voice, video and companionship for the greater good.**

Keywords: Computers, companionship, company, audiovisual, voice, video, face, addiction, Covid, ethics, training, therapy, privacy, data, simulate.

I. INTRODUCTION

Computers are being used for computation, entertainment, video games, communication, the internet, robotics, virtual reality, etc. But now, there is also a looming new possibility of the use of computers for human-like audio-visual interactive companionship. This may even be the next breakthrough or revolution in computer technology.

The advances in computer generated voices have resulted in the now-familiar lady of GPS giving driving directions on the road. The interactive voice of *Alexa* (by Amazon) in our homes responds to our voice by searching the internet, playing music, etc. It is now possible to make human voices and words come from a computer generated moving known face on the screen.

Today's computer voices have taken on a simplistic personality, introducing the possibility of extending the personality to provide companionship. The challenge lies in copying the intricacies of the human voice and thinking, and supporting it with a simulated video of a talking face.

A. History of Computers

On an approximate time scale, the role of computers evolved from simple computation in the 1950s, to entertainment and video games in the 1980s, to data communication and the internet in the 1990s, to robotics and virtual reality in the 2000s.

New generations of robots were created which moved smartly and spoke intelligently. But people were often more interested in what the robot was saying than how they looked and moved. People were largely interested in their voice, her words, and her personality.

In the field of virtual reality and human entertainment, the aim in the past may have been to create an artificial attractive partner with human-like movements. There may also be a large demand for companionship and non-physical interaction with a partner, such as with a computer-generated voice. The need may be less for attractive 3D graphics, and more for a pleasing voice and personality. For people in their 40s and 50s, their needs may shift to just plain human-like companionship with a voice and face.

B. Findings of this Study

With the real possibility of audiovisual companionship with computers, the question arises what does the future hold? How can we promote audiovisual companionship to be beneficial, rather than detrimental?

Market forces may introduce computer voices with less robot-like and more human responses. The personality, gender, age, and experience of the computer voice and image may be chosen by the user. Does the user want in his home, the voice and simulated video of John. F. Kennedy, or Marilyn Monroe, or Joe DiMaggio, or Michael Jackson? Or does the user want the voice and video of her husband or child, who spends long hours away from home? More ominously, does the user want the simulated company of a deceased relative, like a parent, or a grandparent?

In these times of Covid, *Zoom* and *Google Meet* have become the new norm for meeting colleagues, relatives and friends. A simulated voice and image of a person will appear even more to be the reflection of a real person on the other

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

side.

There may be entertainment for the user with a simulated voice and face while watching a TV game. There can be therapeutic companionship, such as at a hospital or an old-age home. A simulated talking face in monitors around the home or in the hospital room would complement the illusion of a companion.

Would a simulated companion be beneficial to the user, or would it become harmful or even addictive? To the terminally ill, dying of say Cancer, can computer companionship be like medical marijuana, to ease the pain and suffering before death?

While using one's energy and resources on a simulated voice and face may be counterproductive, it may also act as training for future real encounters.

Could all these interactions and information from the user be used unethically by the company? Will the government invade the user's privacy by accessing his/her private conversations with the computer? Will law enforcement use the data for advance knowledge of a crime?

This paper explores the possibilities and ethics of audiovisual companionship in the coming years.

II. Literature Review

The concept of computers for companionship is relatively new, as is evidenced by the small amount of literature on this subject. The debate and discussion so far has been when (and if) computers will reach the capability of the human mind (singularity). But this debate has distracted from the very real possibility of companionship with computers.

It has been proposed that computers respond to human emotion [1, 2] according to the needs of the user. It was suggested that there are inherent dangers of humans bonding with social robots [3, 4]. Computers have been likened to the human spirit [5]. There has been research on children interacting with computers both as a learning tool [6] and as therapy for autism [7]. Computers have taken on the personality of pets [8] There has been discussion of human companionship with smart-home devices [9,10]. This paper extends the concept of companionship much further.

III. Spreading Isolation and Solitude

Today's technologically and economically successful society has a negative side, which is the spread of isolation. Covid has confined people in their homes, compounding the isolation. With towns and cities becoming more spread-out and dependent on cars (urban sprawl), there is less face to face interaction between people.

Loneliness is a problem both in developed countries like the US and in developing countries like Bangladesh. The percentage of one-person households in Canada has increased over 1951 to 2016 from 7 % to 28 % (Statscan). In the US, the number of one person households has increased over 1960 - 2017 from 7 million to 35 million (Statista).

Loneliness is often a problem for the elderly, whether living alone or at a retirement home. Children may move to college, leaving parents by themselves, contributing to the empty nest syndrome. Residents of old age homes are known to be lonely. The sick often become lonely, whether at home, in a hospital, or an old-age home.

In *Talk Radio* on the FM radio channels, speakers talk to each other on a contemporary topic. People living alone may leave this Talk radio on all the time, to create a sense they are not alone at home.

Those living alone may tend to talk to themselves, at first in their minds, and then in a low audible voice.

B. Companionship with Inanimate Objects

Perhaps the best examples of companionship with inanimate objects are children with their teddy bears and stuffed toys. In the movie, *Cast Away*, Tom Hanks forms a human-like bond with a crude face painted on a coconut.

Computer simulations have created pets (popular in Japan), that require regular feeding and attention and even simulate death.

IV. Existing Personality of The Computer Voice

The issue is that *Alexa* already has a somewhat robot-like personality. *Google Assistant* will ask the user how he is doing. It will even suggest giving a nickname to the user, such as “Chatterbox.”

It is socially unacceptable to use abusive language with the computer voice. A mother may reprimand her teenage son for speaking inappropriately with *Alexa*. Children may be expected to behave as if *Alexa* or *Google assistant* were real people, effectively helping in the upbringing of children.

The author's sister once asked *Alexa*, "Would you go out with my brother?," to which Alexa replied, "I don't know about that," giving the author a mild feeling of rejection(!).

These responses of the voice give it a rudimentary personality. Grown ups are expected to use socially accepted language (no abusive words or language) with the voice, making it a rudimentary training ground for polishing the social skills of adults.

Today’s computer voices (*Alexa* and *Google Assistant*) basically know everything that can be searched on the internet. The voice knows everything, unlike a real human companion.

Today’s voice hardly has any background, like a place and date of birth, primary, secondary and high school, university.

Social conversations with the computer voice usually do not go on for long periods of time. This could be because long interactions would take up bandwidth and computer time of the company, with little value addition for them. The solution may be to have companionship be generated at a home computer, rather than at a far-away server of the company.

Interactions with the computer voice create a little company at home. It is better to talk to the computer voice

than to talk to oneself.

V. Overcoming Monotony of Computer Voice

What makes computer voice monotonous, and mainly a repository of information? The monotony could be because the voice has a single tone and volume, is designed to find facts, is repetitive, and does not remember past interactions. So what can be done to overcome the monotony, making the voice suitable for companionship?

A. Tones and Volume of the Computer Voice

We can start by programming more tones of voice into the computer. The voice may sound happy, sad, impatient, etc.. The voice may express mild impatience if the user asks the same question repeatedly. The voice may sound happy if the user is able to ask and talk intelligently, saying, "Yes, you are right!"

The voice may be sleepy or drowsy, saying, "I think I will get some rest now."

The voice may be hurried, "Please hurry. Or else you will be late for the meeting"

The uniform volume of the voice only adds to its monotony. By altering tone and volume, there can be the illusion of a companion speaking from another room. Or the volume and tone can simulate a companion speaking barely audibly in the same room, or even lying in the same bed.

B. Remembering Data on User

The computer voice may remember the user's family and education, traits, life events, etc. From the user's date of birth, the voice can wish him/her happy birthday. The conversation can be steered towards interests matching the user's age. A teenage user may be directed to his/her interests in fashion or social events in the downtown area.

An older user may be reminded of past events from say the 1960s, such as, "Jenny, remember Woodstock? That's when we had great music."

The companion may remember the user's schools, universities, and workplaces. It can ask, "Ned, didn't your school have a good football team?"

Workplace information mentioned in a CV can be used for a statement, "Hey, your old company Schlumberger's stock is rising fast."

The companion can use knowledge on parents, siblings, spouse and children of the user. "Judy, it is your mom's birthday tomorrow. Why don't you send her flowers?"

C. Remembering Past Interactions with the User

Perhaps most important, and easily implementable would be a memory of past discussions with the user. Example would be the voice saying, "Harriet, you asked me that before." Or, "I played John Denver yesterday. Did you like my selection?"

VI. Improving Personality of the Voice

Today's voice may remind the user, "You have an appointment with your uncle at 7:00 pm,"

But the voice may have a more engaging personality by greeting the user, "Good morning. How are you today?"

Or the voice may respond to a greeting, "I am fine. And how are you?"

Google Assistant already greets and conducts chit chat, and the challenge may be to extend the chit-chat. The voice may ask questions like or initiate an interaction by, "How is your cold today?" Or, "What are you doing in the kitchen?"

"Did you solve the problem at the office?" or "Did you talk to your boss?"

In presence of bad news, the computer voice can say, "I am sorry your friend came down with Covid."

Detecting the crying sounds of a user, the voice may say, "Maybe we should be with friends at this difficult time."

A. Length of response

To overcome monotony, the computer-generated voice may give answers ranging from long to short. A short answer may be "Yes" or "No."

A medium length of response from the computer voice may be, "Yes, I had a good day looking up facts on the internet."

A lengthy response from the companion may be, "Let me tell you what happened in the game today. The quarterback John Bower started well, but could not match the opposing players,..."

B. A Sense of Humor

Perhaps hardest to program would be a sense of humor. Initially, the computer may be programmed to laugh with the user. It would not be difficult to program the computer to laugh at adult oriented comments by the user, such as:

"The news caster looks good. I wonder whether he/she works out." Or, "I would marry him/her if given the chance."

The voice may respond with a short laugh, or a comment like, "Ha ha. He/she looks muscular. You can become his facebook friend. "

VII. Additional Sensors and Graphics

There may be additional sensors, such as motion detectors in the rooms of the house. That would allow the voice to ask questions like, "What are you doing in the kitchen?" or, "Shouldn't you fix the car today?"

If the user is spending too much time on the couch or bed, the companion voice can say, "Perhaps, we should go out for a walk."

A monitor can show an image of a person lying on a bed, or a person on a monitor while watching a game.

In hazardous situations, such as in a laboratory, there is often a "buddy rule" that one should not work alone. A companion can help in an emergency, and also keep a person mentally grounded. A solitary person may become lost in

thought, making him more likely to make mistakes. A computer-generated voice in a lab may keep a user alert and safe.

C. Accompaniment While Watching Television

The voice and supporting video may accompany the user in front of the television, commenting on a football game or an ongoing TV program.

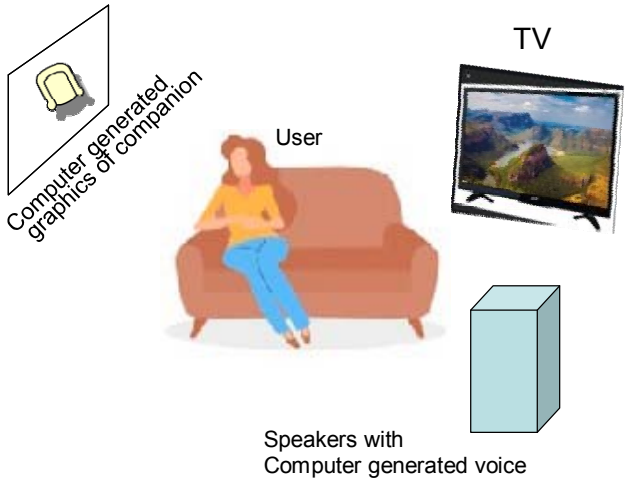


Figure 1. Computer generated voice and supporting face while watching TV

VIII. Limited Knowledge, with Some Education and Experience.

The ability to access the internet and extensive knowledge of the computer voice can be perceived as unreal. So there can be an artificially imposed limit on what the computer voice knows and can talk about.

The background and education of the voice can be according to the needs of the user. It could have gone to some named school, college, and university. It could have had extensive work experience, when the human user is an older person.

A voice and personality used for training social interaction to children need not know quantum physics. Such a voice can be like a mother; helpful, and yet authoritative.

A highly educated person should be able to choose a simulated companion with knowledge on his own area of expertise, like General relativity.

The computer voice at an M.S. or PhD thesis defense may be able to flag plagiarism and inconsistencies (with internet), in the verbal presentation of the student.

D. Multiple Personalities

The user may be able to choose from multiple personalities, such as by saying the following.

"I grew tired talking to middle-aged John yesterday. Can I speak to the younger Jane now?"

To have a sense of continuity, the user may not easily be able to alter the personality of the computer voice. For continuity, the age, gender and background of the computer voice may not be readily changed by the user. Just as the computer pets of Japan die, there can be an irreversible simulated death of a simulated companion.

Like in Talk TV, there can be a choice of channels, discussants and topics. The computer voices could be talking to each other, based on the published news, and other sources of information. "John, what is the news with Covid around the world?"

"Jim, it has not been getting better. In some countries there is a second wave. Looks like restrictions on international flights will not be going away anytime soon."

IX. Personality Types - Actual Vs. Synthetic, Living vs. Deceased

The voice may be a simulated real or an imaginary person. The voice could be a famous personality, such as a scientist, a statesman, an actor/actress or a sportsman. If voice recordings are available (such as for John F. Kennedy), today's technology allows synthesis of new words and sentences based on past audio recordings. For a person deceased for a century or more, such as for Newton, Abraham Lincoln, or George Washington, the voice can be synthesized from existing recordings of others.

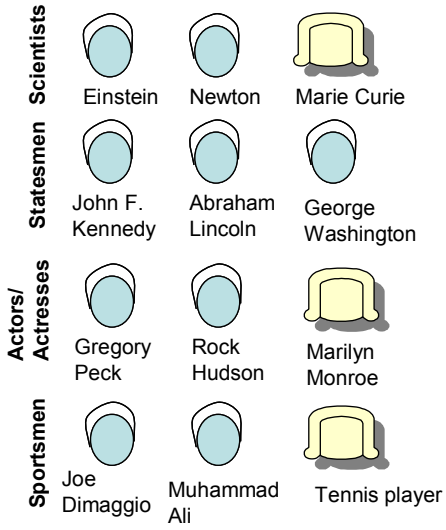


Figure 2. Famous personalities as choices for companions

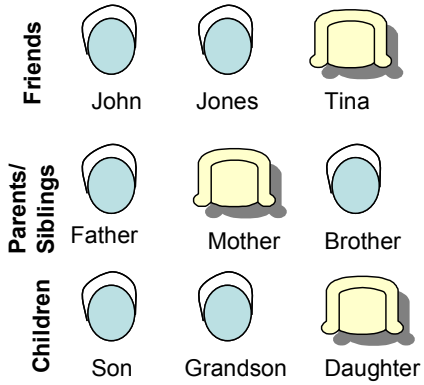


Figure 3. Friends and relatives as choices for companions

X THE ETHICS OF COMPANIONSHIP WITH COMPUTERS

There exists the dark possibility that the user may bond too closely with the computer. Should a companion talk like a relative, such as a child, parent, or spouse, or sibling? The companion could fill-in for the relative during their absence. The companion could address the user as mom, dad, honey, etc., with questionable effects on the mind. This could fool the mind into thinking there is interaction with the relative or friend, when actually there has been none.

There is the even darker situation of the companion remaining behind after the death of the person. What about a mother who has had a miscarriage, interacting with, what to her is the deceased child's computer voice, complete with a growing and changing personality?

Would computer companionship be the new addiction, after the smart phone, *Facebook,* and *Whatsapp*?

Regardless of the ethical issues, it is expected that market forces will prevail, and companionship with computers may become a breakthrough.

XI. DATA AVAILABLE TO COMPANY AND GOVT.

In human interactions with real people, there is the underlying assumption of privacy, meaning the discussions and interactions will not be shared with others. Unfortunately, the interactions with a computer-generated companion will not enjoy this privacy, and will be accessible to the company and ultimately to the government.

The company or the government can use such computer companions to influence the thoughts of the user. Today's voice interactions are already being used to advertise products to the user. The voice of tomorrow can advertise products like, "You should check out the new iPhone model."

Law enforcement can profile and even arrest users based on their confidential conversations with the voice. For example, a computer voice of a child can be the ears of law-enforcement, listening in on whether the mother had an illegal abortion.

The so-called Big-brother of novel, *1984* by George Orwell, may take the form of this computer-generated companion. Taking the voice of a baby or child, the companion may lull the user into surrendering their innermost thoughts. The Orwellian saying would become, "Big Brother is watching, *and talking* with you."

Through the voice companion, the government can influence the user to become a law-abiding and hard-working person, taking care of his family and paying taxes. "John, I think you should go to sleep. You have an early day tomorrow." Or, "Jim, its almost the middle of the year. Are you preparing your taxes?"

"Jane, you are listening to radical radio again. How much longer will we listen to all these conspiracy-theorists?"

XII. CONCLUSION

With today's technology of voice and supporting facial video generation, there is a likelihood that companionship with a simulated voice and face may become the next new computer breakthrough. Covid has only served to confine people to their homes and increase social isolation. Urban sprawl is making towns and cities more spread out. Users are spending more time with their smart phones and computers, increasing real-life isolation and solitude.

This brings the possibility of extending the rudimentary personality of computer voices further to make them much more suitable as companions.

The right tone and volume (and supporting video) can indicate the simulated companion is speaking from another room, or lying right next to the user.

This companionship can be for enjoyment, such as when watching a game, or for therapeutic purposes, such as at an old-age home or a hospital.

Both children and adults are expected to speak politely to the computer voice, effectively acting as training ground for enhancing the social skills of children and adults.

The voice can be used for training parents how to raise a child, or for training a teenager how to interact with the opposite gender. The voice could simulate a patient for training doctors and nurses.

The Buddy rule in the lab means a person should not work alone. The buddy may be a computer-generated companion helping to keep the user alert.

Ethical questions arise about recreating the voices of living or deceased persons. There is the danger of the user bonding too closely to the computer voice and face.

Data on interactions will be available to the company and to the government, infringing on the user's privacy. The company may use the data to advertise products to the user.

Law-enforcement may use conversations to profile or even arrest a user. A child-like voice can be the ears of law enforcement looking for evidence of a crime. Can a woman be arrested for telling a computer voice that she has had an illegal abortion?

The government may shape the personality of the user, inspiring them to be law-abiding, hard-working, tax-paying citizens. George Orwell's quote will become, "Big Brother is watching, and *Talking* with you."

Will computer companionship be the new addiction after the smartphone? Can simulated companionship act like medical marijuana, to ease the pain and suffering before impending death?

Regardless of the ethical issues, it is expected that market forces will prevail, and computer-generated companionship will become the new normal in some years. Regardless of whether we like it or not, our society must find ways to regulate the computer generated companionship for the greater good.

REFERENCES

[1] R. W Picard, Jonathan Klein; Computers that recognise and respond to user emotion: theoretical and practical implications," *Interacting with Computers*, Volume 14, Issue 2, Pages 141–169, 1 February 2002,

[2] Rosalind W Picard ; Jonathan Klein, "Computers that recognise and

respond to user emotion: theoretical and practical implications,: *Interacting with Computers* (Volume: 14 , Issue: 2 , Feb. 2002), pp. 141 - 169

[3] Patrick Lin; Keith Abney; George A. Bekey, "The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots," in Robot Ethics: *The Ethical and Social Implications of Robotics*, MITP, 2012, pp.205-221

[4] M Scheutz *Robot ethics: the ethical and social implications of robotics*, 2011, books.google.com

[5] Turkle S. *The second self: Computers and the human spirit*. Mit Press; Sept. 30, 2005.

[6] Muller AA, Perlmutter M. "Preschool children's problem-solving interactions at computers and jigsaw puzzles." *Journal of Applied Developmental Psychology*. 1985 Apr 1;6(2-3):173-86.

[7] Dinah K.C. Murray, "Autism and information technology: therapy with computers," Chapter 7, *Autism and Learning*, Taylor and Francis Group, 1970.

[8] Ding-Bang Luh, Elena Carolina Li, Yu-Jung Kao; "The Development of a Companionship Scale for Artificial Pets," *Interacting with Computers*, Volume 27, Issue 2, 1 March 2015, Pages 189–201, https://doi.org/10.1093/iwc/iwt055

[9] Byounggwan Lee, Ohkyun Kwon, Inseong Leeb, Jinwoo Kim, "Companionship with smart home devices: The impact of social connectedness and interaction types on perceived social support and companionship in smart homes", *Computers in Human Behavior*, Volume 75, October 2017, Pages 922-934,

[10] Turkle S. "A nascent robotics culture: New complicities for companionship." *American Association for Artificial Intelligence Technical Report* Series AAAI. July 2006.

Dual Band Monopole Antenna For WLAN MIMO Applications at 2.4 and 5 GHz

Gokul Nath S J
Dept.of of Electrical Engineering
Indian Institute of Technology
Palakkad, India
121801015@smail.iitpkd.ac.in

Aswin Raj K
Dept.of of Electrical Engineering
Indian Institute of Technology
Palakkad, India
121801015@smail.iitpkd.ac.in

Soumik Dey
Dept.of of Electrical Engineering
Indian Institute of Technology
Palakkad, India
soumikdey1990@gmail.com

Mohammad Abdul Shukoor
Dept.of of Electrical Engineering
Indian Institute of Technology
Palakkad, India
abdulshukoor872@gmail.com

Athul O Asok
Dept.of of Electrical Engineering
Indian Institute of Technology
Palakkad, India
athulasok14@gmail.com

Sukomal Dey
Dept.of of Electrical Engineering
Indian Institute of Technology
Palakkad, India
sdey28@iitpkd.ac.in

***Abstract*— A dual band monopole resonator antenna is proposed in this paper. The antenna works at two different bands with central frequencies of 2.4 GHz and 5 GHz. To cover the two bands of WLAN (i.e., lower band of 2.4 GHz and upper bands of 5 GHz), two monopoles are used: one working for the lower band and the other working for the higher band. However, there is a limit in most of these antennas to obtain the wideband characteristic especially at the 5-GHz band for WLAN applications, when operating simultaneously at 2.4 GHz. The proposed antenna achieves a good bandwidth at the upper frequency, the data rate has been increased with the help of MIMO configuration. The simulated results are giving a very good return loss of -22.14 dB for 2.4 GHz band and -35.43 dB for 5.0 GHz for the single antenna.**

Keywords—Dual band, monopole antenna, multiple input multiple output.

I. INTRODUCTION

With the increase in need for high data rate in wireless communications making substantial demands on bandwidth and Interoperability within the physical layer, which has resulted in wireless devices expected to function in different multi-band environments. To serve this purpose, a multi-band antenna module with a low profile and miniature size has attracted considerable attention. In addition, the antenna shall have the features of effective impedance matching over broadband, an acceptable gain, and consistent radiation patterns at different frequency bands. Among different types of antennas, monopole or dipole antenna can easily fit for the above-mentioned objectives. Monopole antennas is preferred over the dipole antenna due to the ease of feeding and space constraints. The bandwidth of the monopole or dipole antenna can be enlarged by widening the width of its resonating arm without substantial changes in the radiation pattern over the operating band [1]. The increase in arm width results in a change of the current distribution from the conventional sinusoidal pattern, which helps in broadband impedance matching. The total length of the monopole antenna is generally quarter wavelength long and has symmetric current distribution in the two arms of the monopole with respect to the plane of electrical symmetry [2]. Printed monopole antenna enjoys the features of low fabrication cost, lightweight, easy integration of monolithic microwave integrated circuit (MMIC) components, and simple feeding technique. In recent years different printed monopole antennas for multiband application have reported

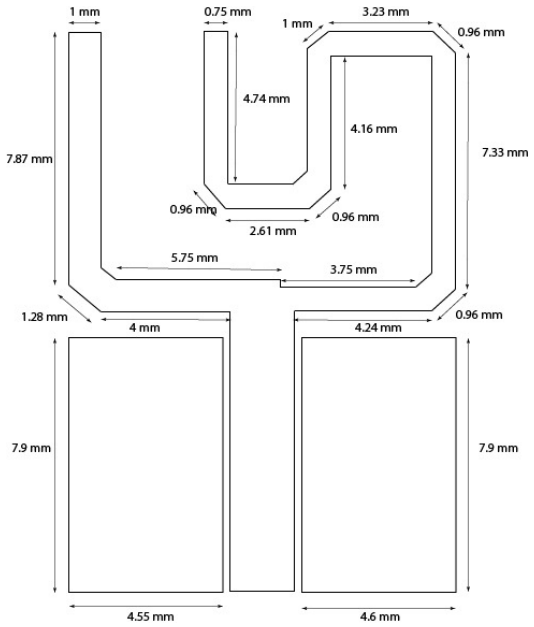


Fig. 1. Prototype configuration of the designed dual band monopole.

[3]–[6]. In [4] coplanar waveguide (CPW) fed planar inverted F antenna (PIFA) to operate at GSM and worldwide interoperability for microwave access (WiMax) band. Zeroth order resonance of dual composite right/left-handed (DCRLH) transmission line structure has used to get multiple frequency bands for monopole antenna in [5].

The proposed antenna combines two folded strips and is fed by a CPW transmission line, which radiates at 2.4 and 5.0 GHz. The growing demand for WLAN technology has led to an interest in integrating the 2.4-GHz and 5-GHz frequency bands into a single device that requires a dual-band antenna. The WLAN antenna not only has a wider frequency band, smaller size, and is easier to install, but it also has a high radiation efficiency [7]. MIMO wireless technology can highly improve the data rate, capacity, and link reliability of wireless systems through multi-path data transmission and reception. MIMO system is currently employed in 4G user equipment and is a promising technology for use in the future 5G mobile terminals [8]–[9]. The last section of the paper is focusing on implementation of the MIMO for the current design.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

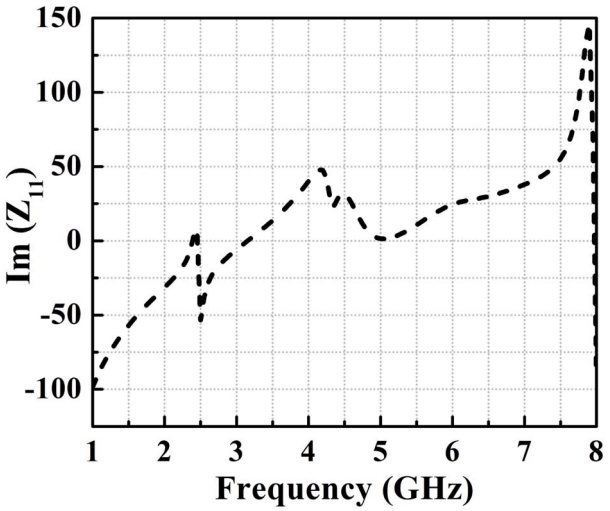


Fig. 2. Imaginary part of the input impedance for the designed dual band monopole antenna.

II. Dual Band Monopole Antenna

A. Antenna Geometry

The proposed antenna is constructed on FR4 substrate of 18 mm × 14 mm. Substrate with thickness of h=1.6 mm and relative dielectric constant of 4.4. Fig. 1 shows the configuration of the antenna which consists of two branches resonating at 2.4 GHz and 5 GHz. It consists of a coplanar ground plane, a radiation structure, and a feed line. A CPW transmission line of 50 Ω, which consists of a signal strip width of 2 mm and a gap distance of 0.45 mm and 0.4 (right and left respectively) between the single strip and the co planar ground plane, is used for feeding the antenna. The width has been further optimized to get a better performance. The width of the microstrip feed line is fixed at 0.75 mm and 1mm for 2.4Ghz and 5.0GHz respectively. Dimensions of the port has been found out using the equations (1) and (2),

$$W = 3*(g_1 + g_2 + W_f) \quad (1)$$

$$H=6*h \quad (2)$$

where g_1, g_2 are the gap lengths between the co planar ground plane and the feed line, W_f is the width of the feed line, h represents the height of the substrate. The dimensions of the co planar ground plane is 7.9 mm × 4.55 mm and 7.9 mm × 6 mm. The approximate dimensions of the feed line is optimized so as to get an impedance matching. The two monopoles on the left and the right will resonate at 5 GHz and 2.4 GHz respectively. The length were chosen as λ/2 for each resonating frequency.

From Fig. 2 it can be seen that the proposed antenna is an open circuited parallel resonating circuit which will resonates at even multiples of λ/2. We have optimized the length after taking the $\lambda_{eff}/2$ as a reference.

$$\lambda_{eff} = \frac{c}{f_o\sqrt{\varepsilon_{eff}}} \quad (3)$$

This can be verified by looking at the imaginary part of the impedance. From the Fig. 2. It can be observed that at

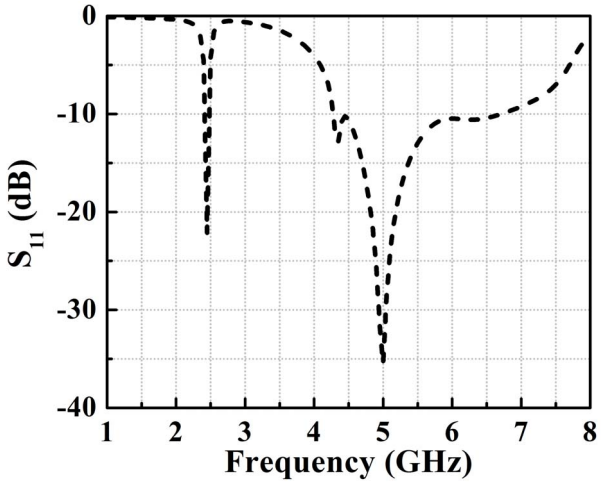


Fig. 3. Simulated reflection coefficient of the dual band antenna.

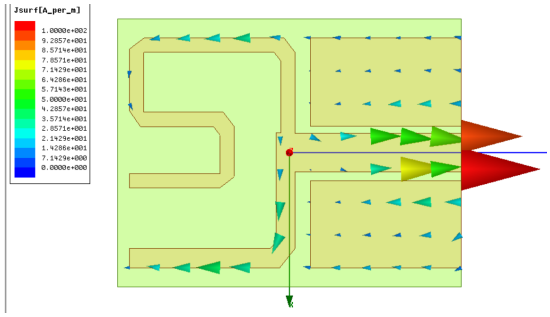
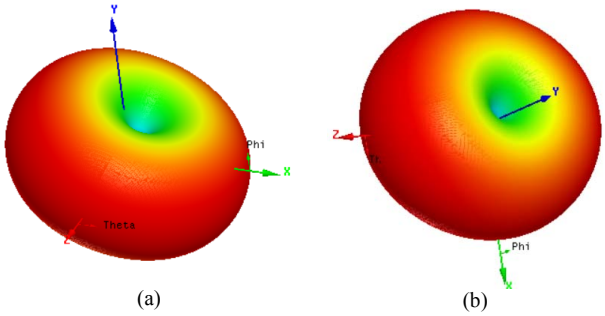


Fig. 4. Simulated surface current distribution at 5 GHz.

frequencies 2.4 GHz and 5 GHz, the circuit is acting like an open circuited parallel resonator. Here The impedance abruptly changes its value from maximum positive to the most negative value. We can exploit this for understanding to our dual band antenna design. The gap length between the feed line and ground plane will affect the higher band frequency. While observing the current distribution, we could observe that the direction is opposite in both the surfaces. This is an indication of the capacitive effect between the feed line and the ground plane.

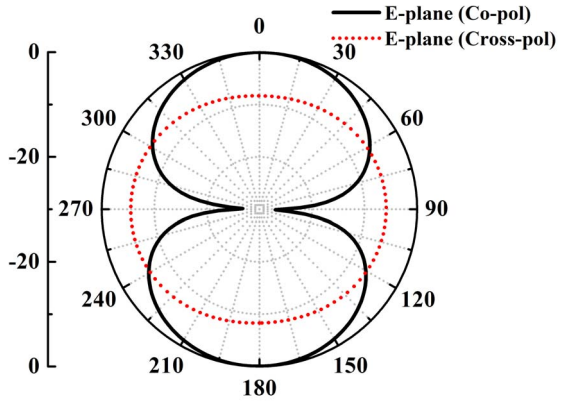
B. Simulation Results

The CPW-fed monopole antennas with various parameters were simulated and studied to demonstrate the proposed results. The simulated results are obtained using the Ansoft simulation software high-frequency structure simulator (HFSS) [10]. Fig. 3 shows the simulated return loss for the dual band antenna. The designed monopole antenna is showing a dual band resonance with S_{11} of -22.14 dB for 2.4 GHz and -35.43 dB for the frequency 5 GHz. It is observed that with the increase in gap length, the lower band frequency almost remains same while the upper band frequency is affected. Observing the surface current distribution in the feed line and the ground plane, it is seen that a capacitive effective is acting between them as seen from the Fig. 4. This is shifting the upper band frequency when the gap length is changed. The triangular cut made at the edges of the strip is used to increase the current distribution through the monopole. The strips were cut at the edges, where there will be a hindrance to the current flow.

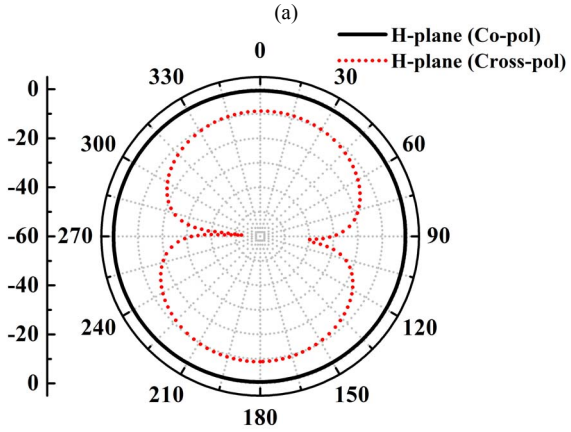


(a) (b)

Fig. 5. 3D radiation patterns of the monopole antenna at frequencies (a) 2.4 GHz and (b) 5 GHz.



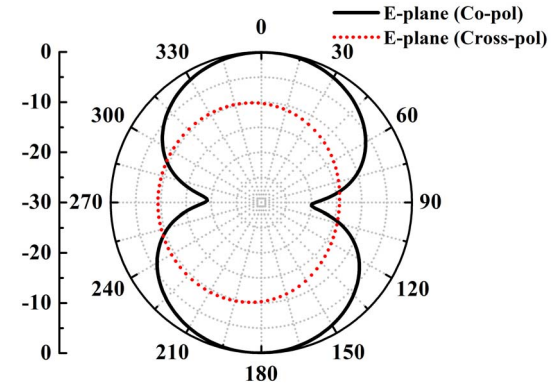
(a)



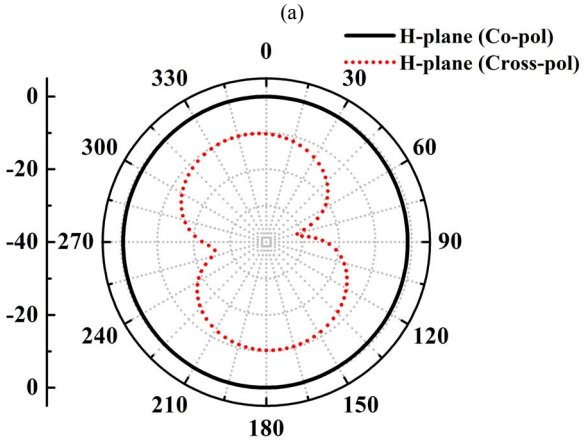
(b)

Fig. 6. Simulated radiation pattern of the monopole antenna at 2.4 GHz (a) E-plane (b) H-plane.

This has also resulted in a good impedance matching between the antenna at both the frequency. The upper frequency band is showing a very large bandwidth of about 2.425 GHz while the lower 2.4 GHz is primarily a narrow band. The 3-D radiation patterns are shown in the Fig. 5 at the two different frequencies. It is observed that the radiation pattern is approximately omnidirectional especially for the higher frequency. The measured peak gains for the proposed antenna is around 2.445 dBi at 5 GHz. The co-pol and cross-pol radiation patterns of antenna are for the two frequencies 2.4 GHz and 5 GHz are shown in Fig. 6 and Fig. 7 respectively. It is clearly see that the co-pol and cross-pol difference is less than -10 dB which indicates that a good amount of power is utilized in for a particular polarization results in the plane the corresponding plane.



(a)



(b)

Fig. 7. Simulated radiation pattern of the monopole antenna at 5 GHz (a) E-plane (b) H-plane.

Fig. 8. Schematic for the 2 × 1 dual band MIMO antenna.

III. Geometry of Mimo Antenna

The MIMO implementation is done to increase the data rate of the entire system. A 2 × 1 MIMO has been implemented using the same design that is described above. The antenna axis has been placed perpendicular to each other to reduce the mutual coupling between the two antennas. The distance between the antennas have been optimized to get the required results making λ/2 as the reference distance. Fig. 8 depicts the geometry of the 2 × 1 MIMO configuration. Two dual band monopole antennas have been placed with orthogonal feeding position to improve the

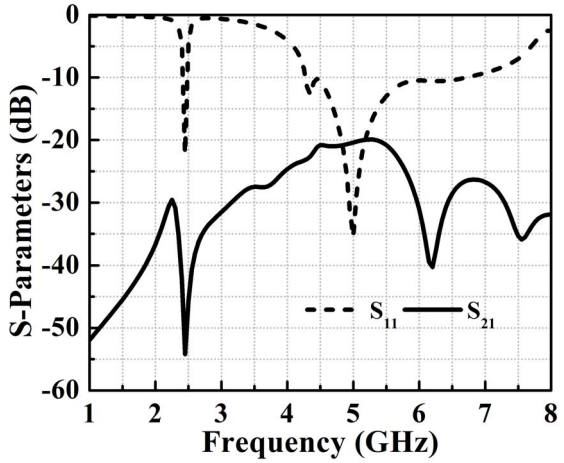


Fig. 9. Simulated return loss and isolation between the two antennas.

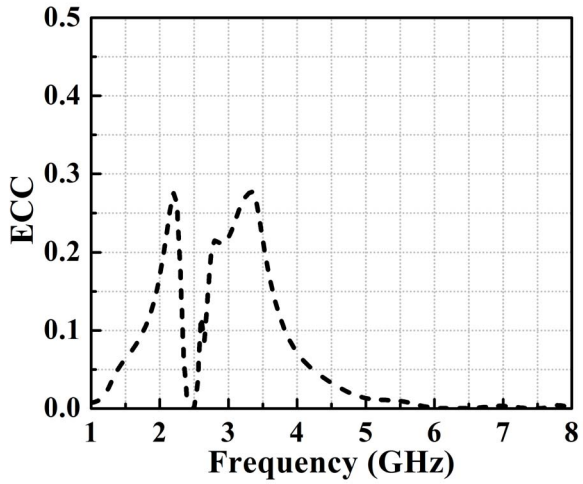


Fig. 10. Simulated ECC from S-parameter for MIMO antenna.

isolation between them. The simulated results for various S parameters have been studied and analyzed using the HFSS software. The Fig.9 shows the results of the two bands at 2.4 GHz and 5 GHz. The S_{21} results are also plotted vs frequency. The maximum value of S_{21} is around -20 dB indicating a reduced mutual coupling effect between the antennas.

The Envelope Correlation Coefficient (ECC) of the MIMO antenna is also studied. The ECC is calculated by using the following expression,

$$ECC = \frac{\left|S_{11}^{*}S_{12}+S_{21}^{*}S_{22}\right|^{2}}{\left(1-\left|S_{11}\right|^{2}-\left|S_{21}\right|^{2}\right)\left(1-\left|S_{22}\right|^{2}-\left|S_{12}\right|^{2}\right)} \quad (4)$$

The expression is fed into the software and the plots were made as presented in Fig. 10. ECC is well below 0.5 for both operating frequency which suggest the possible application of the designed configuration for WLAN application.

IV. Conclusion

The proposed antenna has a simple configuration and is easy to fabricate. To obtain the required impedance matching, the gap widths have been optimized by parametric analysis. The designed antenna satisfies the -10 dB return loss requirement
at 2.4 GHz to 5.0 GHz band and also provides a good radiation pattern at the higher frequencies. Further the proposed antenna array also shows superior performance robustness and MIMO capabilities under different hand-grip conditions. In conclusion, the proposed MIMO antenna design and array is promising for WLAN applications.

Acknowledgment

The authors would like to express their profound gratitude to the Director of Indian Institute of Technology Palakkad for setting up Central Instrumentational Facilities. This work is supported by Science and Engineering Research Board, Govt. of India under the project no: ECR/2018/002258.

References

[1] C. Balanis, Antenna Theory, Analysis, and Design, 3rd ed. New York: Wiley, 1997.

[2] A. T. Mobashsher and A. Abbosh, "Utilizing Symmetry of Planar Ultra-Wideband Antennas for Size Reduction and Enhanced Performance," *IEEE Antennas Propag. Mag.*, vol. 57, no. 2, pp. 153-166, Apr. 2015.

[3] M. Wu and M. Chuang, "Multibroadband Slotted Bow-Tie Monopole Antenna," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 887-890, Dec. 2014.

[4] A. Soliman, D. Elsheakh, E. Abdallah and H. El-Hennawy, "Multiband Printed Metamaterial Inverted-F Antenna (IFA) for USB Applications," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 297-300, Sep. 2014.

[5] H. Li, G. Wang, X. Gao and L. Zhu, "CPW-Fed Multiband Monopole Antenna Loaded With DCRLH-TL Unit Cell," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 1243-1246, Feb. 2015.

[6] M. Naser-Moghadasi, R. A. Sadeghzadeh, M. Fakheri, T. Aribi, T. Sedghi and B. S. Virdee, "Miniature Hook-Shaped Multiband Antenna for Mobile Applications," *IEEE Antennas Wireless Propag. Lett.*, vol. 11, pp. 1096-1099, Aug. 2012.

[7] M. van Rooyen, J. W. Odendaal and J. Joubert, "High-Gain Directional Antenna for WLAN and WiMAX Applications," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, pp. 286-289, May 2016.

[8] Q. Nadeem, A. Kammoun, M. Debbah and M. Alouini, "Design of 5G Full Dimension Massive MIMO Systems," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 726-740, Feb. 2018.

[9] Y. Li, C. Sim, Y. Luo and G. Yang, "Multiband 10-Antenna Array for Sub-6 GHz MIMO Applications in 5-G Smartphones," *IEEE Access*, vol. 6, pp. 28041-28053, May 2018.

[10] Ansoft High Frequency Structure Simulator (HFSS), Ver. 13.0, Ansoft Corporation.

Simulation Framework for Development and Testing of Autonomous Vehicles

Ahmed AbdelHamed
Mechanical Engineering
Kettering University
Flint, USA
abde5186@kettering.edu

Girma Tewolde
Electrical and Computer Engineering
Kettering University
Flint, USA
gtewolde@kettering.edu

Jaerock Kwon
Electrical and Computer Engineering
University of Michigan - Dearborn
Dearborn, USA
jrkwon@umich.edu

***Abstract*— The benefits of Autonomous Driving (AD) and Advanced Driver Assistance Systems (ADAS) include providing critical capabilities that would help to reduce accidents due to driver errors, improve safety in traffic, and enhance driving efficiency. The goal of this paper is to utilize the Robot Operating System (ROS) and Gazebo 3D simulator for realizing a fully immersive software-in-the-loop (SIL) simulation framework, which could be utilized for the virtual design, verification, and validation of AD/ADAS features. It supports designing and implementing a software platform and methodology to validate different features for AD/ADAS with a high degree of realism. The paper will also discuss the configuration of the sensors, such as the camera, LiDAR, and RADAR, that are later used as part of a sensor fusion technique that shows the usability of the developed environment.**

***Keywords*— Simulation Framework, Software-in-the-loop (SIL), Robot Operating System (ROS), Gazebo, Autonomous Driving (AD), Advanced Driver Assistance System (ADAS)**

I. INTRODUCTION

In recent years, research and advancements in the fields of sensors, automation, and intelligent systems have greatly helped the automotive companies to offer products in the form of Advanced Driver Assistance Systems (ADAS), that is becoming a necessary feature for today's vehicle owners. The possibility of having a system that can monitor the vehicle's surroundings and the on-board sensors to maintain safety are becoming necessary features in the newer vehicle models. The autonomous and intelligent systems are not considered a myth anymore, but they are becoming a reality and one of the most vital research areas in the automotive industry these days. Development of a system that is capable of navigating autonomously while being aware of the surroundings, without any human interaction, is a challenging task.

Over the last decade the automotive industry has been actively developing ADAS features for driver comfort and safety, and now the technology is advancing towards the development of fully autonomous cars. Advancements in technologies such as, Light Detection And Ranging (LiDAR), cameras, Radio Detection and Ranging (RADAR), and their fusion help provide a 360-degree surround view of the environment around a car, thus enhancing the ability to perceive the dynamic environment around the vehicle in the best possible way. The amalgamation of multiple powerful sensors with the intelligent algorithms help ensure improved vehicle mobility and enhanced safety of the passengers and vulnerable road users such as cyclists and pedestrians.

Even with many of these technological advancements, there are still a number of challenges that need to be overcome to achieve the full potential of autonomous vehicles for the highest levels of safety, driver/passenger comfort, efficiency, and cost before the autonomous vehicles become available commercially for consumers. Today, many universities and research institutions, major companies, and government organizations are working hard to get a fully functional autonomous vehicle on the road. Although it appears they are getting close to achieve this goal with proof of concept and fleets of test vehicles; for a fully robust autonomous vehicle, there needs to be over a billion miles of autonomous driving experience. Developers need to fine-tune their technology to handle well all kinds of difficult scenarios on the road, which is necessary before they can bring their technology for mass production. This however can be facilitated with the help of simulation, as demonstrated by Waymo [1]. Hence, we see a significant need for a simulator. A simulator can speed up the process of validating the driving technologies and reduce the costs for the testing as well. Based on a study made by KPMG, a professional service company, the utilization of autonomous vehicles for transportation and defense applications is expected to show tremendous growth over the coming decade [2]

For autonomous vehicle simulation environments many research groups utilize the Robot Operating System (ROS) [3] along with Gazebo [4]. ROS is a collection of tools and libraries which helps to simplify the task of building robotic and autonomous vehicle systems. Gazebo is a simulation tool that helps to simulate robots or vehicles along with their sensors in various environments. The advantage of using Gazebo is because it has a topic based publish/subscribe model of communication which goes hand in hand with ROS and makes it easy to create distributed and networked models [5]. The Gazebo Plugins help to build a direct communication interface with ROS thereby being a competent simulation tool for developing and testing robots and autonomous vehicles.

In this paper we present the software-in-the-loop (SIL) simulation framework that we designed and evaluated. The

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

developed vehicle model, named MIR vehicle, allows anyone to build a simulated vehicle model in Gazebo for testing using ROS. The evaluation proved that the proposed simulation framework can be an effective platform of SIL modeling and simulation for autonomous vehicles research and development. This paper is being presented as an improvement over a previous version of the simulation environment [6], with easily adaptable vehicle, sensors, and environment models.

II. Related Work

With the rapid technological advancement in the field of automated driving systems, researchers often use simulation environments for the development lifecycle of their algorithms. Companies are investing vast amounts of their budget to identify effective simulation tools that can be used for their developments. A simulation environment with relatively close to real-life physics engine and solid sensor models are favored over others. Currently, there are lots of open-source simulation platforms that can be utilized for various research applications.

One of the earliest developments in the field of autonomous vehicles testbeds is the CAT Vehicle (Cognitive and Autonomous Test Vehicle [7]. This open-source testbed is built over ROS as middleware and provides several packages for the control application and different simulated sensor models. The testbed gives the possibility of transferring the MATLAB or Simulink control model from simulation to the physical platform through a straightforward workflow. The system model also includes sensors and actuators with configurable parameters for monitoring the vehicle's motion, perception of the objects in the environment and for controlling the driving behavior of the vehicle in a simulated world. The simulation tool also supports logging and playback of data for close analysis of the system's performance [7].

LG Simulator (LGSVL) [8] is a ROS/ROS2 simulator that helps in the testing and development of autonomous driving software systems. The simulator helps as well in speeding up the algorithm development cycle and system integration process. The simulator can publish real-time sensor outputs from sensors including camera, LiDAR, RADAR, GPS, and IMU. The LGSVL is compatible with Apollo and Autoware for easier migration to these frameworks [8].

CARLA [9] is an open-source simulator developed for autonomous vehicles systems research. Built to serve as a modular and flexible API to solve most of the problems involved in the development of autonomous driving system. CARLA is built using the Unreal Engine (UE4.24) to run the simulation and uses the OPENDRIVE maps to define roads and urban settings. The control over the simulation is easily accessible through a Python and C++ API. CARLA simulator consists of scalable client server architecture. The server is controlling the traffic manager, sensors, recorder, and scenario runner. CARLA can be integrated with ROS and Autoware [9].

Another driving simulation software used for the development and testing of autonomous vehicles, ADAS and vehicle dynamics is rFpro [10], which allows the introduction of professional test drivers into the model-based engineering process. The advantage of rFpro is its ability to replicate the vehicle dynamics so developers can evaluate how comfortable, safe, and anxious the passenger feels. The ability of pushing your AI to the limit and testing it in complex situations with real human road users entirely in simulation and nobody would be hurt when something goes wrong, is the best thing these types of systems offer. The co-simulation between rFpro and IPG's Carmaker allow the transfer of sensor data and traffic objects between them [10].

Plenty of other simulation software tools are currently available for commercial use. Major companies are currently investing in the research of building their own simulation platform such as NVIDIA Drive constellation created by NVIDIA, Carcraft created by Waymo, Prescan by Tass international. These and many other simulation software are developed on top of powerful physics engines such as Unity, PhysX, and the Unreal Engine. These all are great tools for testing and validation of autonomous vehicle algorithms. The issue with these tools is the lack of accessibility and cost for many researchers and students with limited budgets. Hence, our goal in this paper is to create a framework which is based on open-source and free software platforms, which are easily accessible to anyone interested, and with the flexibility to create their own environment and vehicle models.

III. Framework Design

The core of this project was to build a simulation framework with an integrated vehicle model and all the necessary sensors that will be utilized to develop and validate autonomous vehicles algorithms. All the procedures that are undertaken were documented to help anyone to replicate the development process of such software-in-the-loop (SIL) environment and the integrated vehicle model. The framework is developed using Python, C++, and point cloud library (PCL 1.8). The simulation environment used is compatible with Gazebo versions 8 and 9, utilizing ROS Kinetic as the framework middleware.

A. ROS

ROS (robot operating system) is a set of packages and libraries that provide the hardware abstraction, sensor drivers, and messaging infrastructure that enable the robot application framework. The reason behind using ROS as middleware is to maintain the modularity of the perceptions and decision-making modules. Installation and environmental setup instructions for the ROS system are explained at the ROS website [11].

B. Gazebo

Gazebo is a 3D simulation engine. It gives the capability to simulate robots accurately by defining the robot joints and links. Joint is a point which is connecting two links, and a link is the actual robot part; it can be an arm, a manipulator, or a wheel. What makes Gazebo a useful tool for building your

environment and simulating your robot on is the robust physics engine that realizes your model with a high degree of fidelity and a decent graphical rendering feature [12]. To integrate the ROS with Gazebo, a set of ROS packages named gazebo_ros_pkgs must be used. These packages provide the interfaces to simulated vehicles in Gazebo to ROS nodes.

IV. ENVIROMENT MODEL

The simulation environment is a modified version of the Open Source Robotics Foundation OSRF (OSRF) car_demo package [13], as seen in Figure 1.

Different road conditions, different lane marking colors and types, traffic signs and traffic lights, and moving as well as stationary pedestrians were added in the environment. An integrated vehicle model was used with a real life vehicle size of Chevrolet Bolt with all the necessary sensors (Camera, LiDAR, and RADAR).

Figure 1: Environment Model

A. *Traffic sign Modeling*

There are several steps that need to be followed in order to create and import a model in Gazebo. For creating a traffic sign in Gazebo, the first step is to create a 3D computer-aided design (CAD) model that assembles the shape of the traffic sign in any CAD software such as SolidWorks, Siemens NX, or Catia. The second step is to add texture on the traffic sign; this step adds the color and the emblem of the traffic sign on the CAD model, this can be quickly done using a free open-source software named Blender using the UV editing tool, the output of which can be seen in Figure .

Figure 2 : Traffic signs in Gazebo

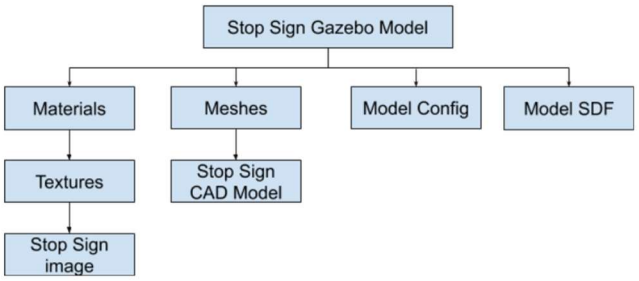


Figure 3. Model Structure in Gazebo

The next step is to export the model into a COLLADA (.dae) file format. After exporting the model to .dae, create a folder named Meshes and paste the .dae file inside this folder, also add the image which was used while attaching the texture on the CAD model inside a folder named Materials > textures, as shown in Figure 3. The steps mentioned here can also be followed for adding texture to your simulated vehicle.

The final step required to complete the traffic sign model is to create the model configuration files for Gazebo. The model.config and model.sdf files are written in xml file format.

B. *Traffic Light, pedestrian, and road conditions Modelling*

The Traffic light, walking/standing pedestrians, and SUV/sedan vehicle models used in this simulation environment are available for importing directly to the Gazebo model. After downloading the necessary models, import them by copying the models to ~/.gazebo/models. After restarting Gazebo, you will be able to find the imported models in the insert tab. The model dimension and model collision properties can be changed from the model.sdf file. The model collision properties are used to prevent the object from being a shadow in the simulation environment. In other words, if the collision properties are not added to the model, and a vehicle is crashing into that model, it will just pass through it; and the object cannot be detected by the sensors such as LiDAR.

There is a plugin named libActorCollisionsPlugin [13] which is for adding collisions to the actor pedestrians. This plugin was used to detect pedestrians using LiDAR pointcloud2 and to prevent the vehicle model from passing through them.

A plugin has been used to control traffic lights [14] from the keyboard through publishing a visual message to the Gazebo topic named /gazebo/default/visual and subscribing to /gazebo/default/keyboard/keypress Gazebo topic.

For moving objects in the Gazebo environment, either a pedestrian or a vehicle, their movement is defined by adding the trajectory in the world file. The method used to move pedestrians is by setting a loop of the set of points the pedestrian moves through within the simulation environment and continuously updating the location.

V. VEHICLE MODEL AND SENSOR PLUGINS

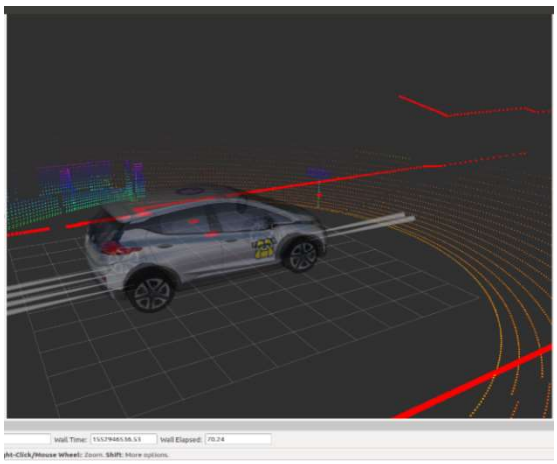


Figure 4. Vehicle Model view in Rviz

A vehicle model with different sensors for autonomous vehicles such as cameras, LiDAR, and RADAR need to be integrated inside Gazebo as seen in Figure 4. The vehicle model is adopted from the OSRF open-source hybrid Prius model [12]. As mentioned before, to import an environment into Gazebo, a .world file needs to be created, which gathers a set of models that can be visualized.

A universal robotic description format (URDF) [16] model needs to be created, the files of which are written in XML format, and it is used in ROS to describe all the elements of a robot and its environment. The vehicle created is matching the real life size vehicle and also the frames orientation and position is matching with the real life research vehicle that has been created for Kettering University's entry in the GM/SAE sponsored AutoDrive Challenge [17].

To be able to navigate the vehicle in Gazebo simulation, a plugin is needed to control the wheels and the vehicle speed. Typically, a PID controller can be implemented to control the vehicle model to maintain the assigned speed. Also, it is possible to control the vehicle using the keyboard by configuring it to respond to specific ASCII codes of the keyboard input. The vehicle controller plugin lets us adjust all the vehicle parameters like the maximum vehicle speed and many more parameters which affect the vehicle performance.

For testing purposes a custom ROS message is created to publish the vehicle readings, such as the vehicle's speed, the transmission shifting status (D, N, R), the throttle percentage, the brake signal as either 0 or 1, and the steering angle from -1 to 1 (1 is the maximum steering value to the left and -1 is the maximum steering value to the right).

A joint state publisher [18] is used to calculate all the transformations between all the defined joints in the URDF with the help of an ignition math package, that provides a library to deal with all the math operations needed for robotics applications [19]. The primary function of a joint state publisher is to manipulate the joints; for instance, it was used to move all wheels with the same velocity. The vehicle description with links and joints in the URDF is shown in Figure 5.

A. Sensor plugins

One of the significant capabilities of self-driving cars is their ability to sense the surrounding environment. Without such ability to detect the area around the vehicle it would be impossible for it to recognize obstacles or other objects of interest. The primary sensors that are commonly used in most autonomous vehicle developments are camera, LiDAR, and RADAR.

There are lots of open-source sensor plugins ready to be integrated with ROS. The main task here will be working on integrating the sensor plugins with your model and adjusting the sensor frames. The sensor frame is one of the essential elements to consider while building your model.

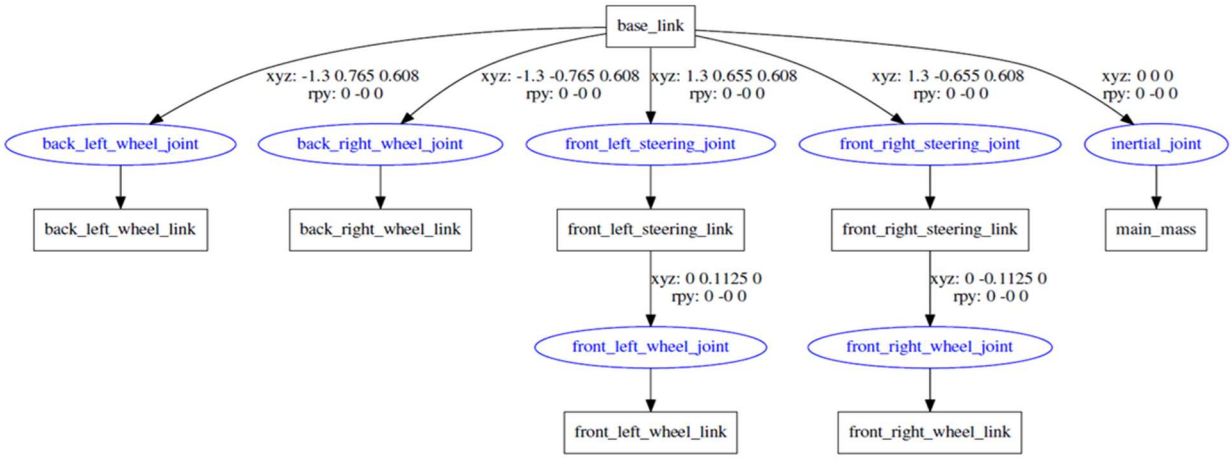


Figure 5. Vehicle Transformation Frames

For clarification, if we detected an obstacle that is 5 meters away from the camera, the transformation between frames will exactly give us how far the object is from any frame in the vehicle. As all the links are linked to the chassis link, therefore by using the transformation ROS package, a relation between them is easy to calculate [20].

1) LiDAR

LiDAR is one of the most commonly used sensors for object detection. We were able to use a Velodyne LiDAR sensor plugin in a pointcloud2 format. There are lots of versions for the Velodyne LiDAR, but for our work here we used the VLP16.

2) Camera

An 800x800 pixel resolution, eight-bit, RGB camera is used as a front center camera, and a frame named camea_link is used to publish the camera sensor messages. The type of noise modeled by the sensor is Gaussian noise, whose parameters such as its mean and standard deviation can be varied in the tag in the sensor link. The libgazebo_ros_camera [20] is part of Gazebo plugins, which is automatically installed when installing Gazebo plugins.

3) RADAR

The RADAR plugin is not generating the same sensor header as the real Delphi SRR/ESR, but it can be easily mapped to publish specific messages, with all the primary sensor data available to be utilized. The idea behind this plugin is that, since Gazebo does have a sensor plugin named the logical camera [21]; this camera plugin can give the coordinates of the object based on the ground truth coordinates of an object in the simulation environment as seen in Figure .

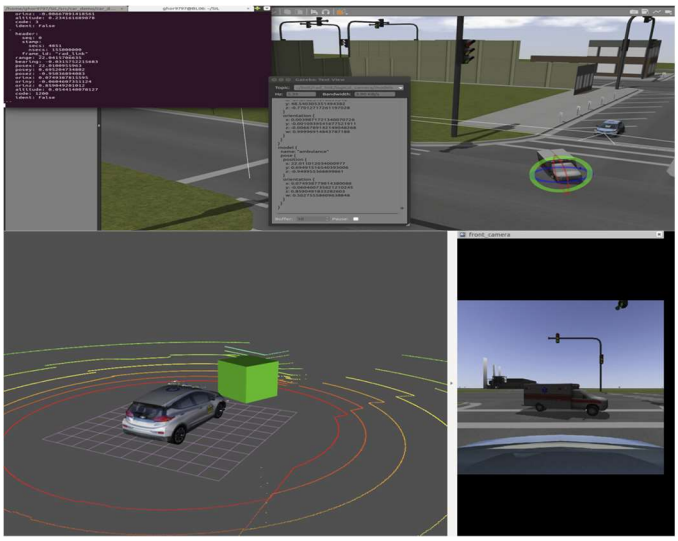


Figure 6 : RADAR sensor view in Rviz

Basically, Gazebo is publishing the coordinates of the objects detected in the simulation environment over a Gazebo topic, the plugin subscribes to that topic and publishes the calculated range, the bearing, and azimuth of the detected object over ROS.

VI. Example Use Cases

a) Data collection for training a neural network. A driver-in-the loop (DIL) system is setup using the developed simulation framework for data collection that can be used as part of development of autonomous vehicle algorithms, such as lane keeping assist systems using neural networks. Since good data helps the network to learn nicely and easily, on the other hand if the data is not proper then however good the network is, it will not be able to learn, predict and converge as expected. The simulation framework presented in this paper was used for data collection task using Logitech G920 dual-motor feedback driving force steering wheel, accompanied with pedals, gear shifter and car seat. Multiple screens were setup to monitor different perspectives of simulation and to observe the output of the image taken by the front mounted camera on the model vehicle [23].

b) Evaluating sensor fusion algorithms. Another application example that proves the validity of the SIL framework is the evaluation of different sensor fusion techniques. The authors developed and tested sensor fusion techniques using camera and LiDAR sensors (see Figure), to demonstrate effective detection of pedestrians and various types of vehicles in the environment. The two types of sensors used in this study are camera and LiDAR, which are commonly used in AD/ADAS applications. Each of these sensors have their own inherent strengths and shortcomings under different operating environments. The two alternative sensor fusion techniques evaluated are: i) data fusion (also referred to as early fusion), and ii) results fusion (also referred to as late fusion). Early fusion techniques fuse the raw sensor readings of the camera and LiDAR to overlay range information from the LiDAR on top of objects detected from the camera vision sesnor. In the results (or late) fusion method, however, separate algorithms process each sensor data and the final determination of the detected object is made based on the detection confidence levels from each of the individual sensors [24].

The complete details of the sensor fusion techniques evaluated along with the performance metrics is being compiled for furture paper presesntation.

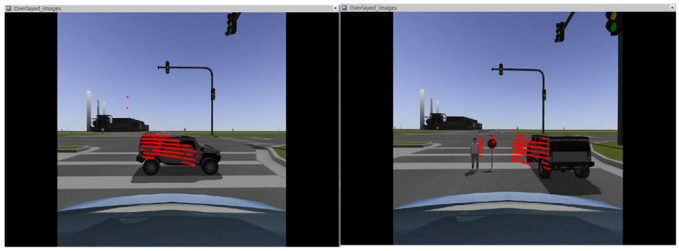


Figure 7: Sensor Fusion testing using the SIL framework

VII. Conclusion and Future Plans

The purpose of this work was to develop a solid vehicle model and simulation environment based on ROS and Gazebo that enables researchers to develop and test various autonomous vehicle algorithms. The developed software-in-the-loop (SIL) simulation framework has been successfully utilized for the development, testing, and evaluation of different sensor fusion algorithms for object detection and classification purposes using camera and LiDAR. The authors are currently preparing another paper to present the sensor fusion techniques and their performance assessment.

In future work, the authors would like to explore controlled testing environments that will allow researchers to add environmental conditions (such as snow, dust, and fog) to the Gazebo model, which will be useful for enhancing the models to represent realistic sensor features.

The presented simulation framework is publicly available and can be cloned from GitHub [25].

Acknowledgment

The authors would like to acknowledge the partial support received from the Kettering University Provost's Research Matching Fund.

References

[1] "Waymo has now driven 10 billion autonomous miles in simulation," *TechCrunch*. http://social.techcrunch.com/2019/07/10/waymo-has-now-driven-10-billion-autonomous-miles-in-simulation/ (accessed Feb. 29, 2020).

[2] "Global Automotive Executive Survey 2019 - KPMG Global," *KPMG*, Feb. 20, 2019. https://home.kpmg/xx/en/home/insights/2019/01/global-automotive-executive-survey-2019.html (accessed Feb. 29, 2020).

[3] M. Quigley *et al.*, *ROS: an open-source Robot Operating System*. 2009.

[4] "Gazebo : Tutorial : Quick Start." http://gazebosim.org/tutorials?tut=quick_start (accessed Jun. 09, 2019).

[5] "osrf / gazebo_models — Bitbucket." https://bitbucket.org/osrf/gazebo_models/src/default/ (accessed Jun. 09, 2019).

[6] M. Ahamed, G. Tewolde, and J. Kwon, "Software-in-the-Loop Modeling and Simulation Framework for Autonomous Vehicles," the 2018 IEEE International Conference on Electro/Information Technology (EIT), pp. 0305–0310 (2018)

[7] R. K. Bhadani, J. Sprinkle, and M. Bunting, "The CAT Vehicle Testbed: A Simulator with Hardware in the Loop for Autonomous Vehicle Applications," *arXiv. 2018. 1804.04347*

[8] L. Simulator, "LGSVL Simulator," *An Autonomous Vehicle Simulator*. https://www.lgsvlsimulator.com/ (accessed Aug. 17, 2020).

[9] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An Open Urban Driving Simulator," in *Conference on Robot Learning*, Oct. 2017, pp. 1–16, Accessed: Apr. 03, 2019. [Online]. Available: http://proceedings.mlr.press/v78/dosovitskiy17a.html.

[10] "ADAS & Autonomous," *rFpro*. http://www.rfpro.com/virtual-test/adas-and-autonomous/ (accessed Aug. 20, 2020).

[11] "kinetic/Installation/Ubuntu - ROS Wiki." http://wiki.ros.org/kinetic/Installation/Ubuntu (accessed Feb. 29, 2020).

[12] "Gazebo : Tutorial : Ubuntu." http://gazebosim.org/tutorials?cat=install&tut=install_ubuntu&ver=9.0 (accessed Feb. 29, 2020).

[13] "Simulated Car Demo." https://www.osrfoundation.org/simulated-car-demo/ (accessed Feb. 24, 2020).

[14] "Example plugin that enables collisions." https://bitbucket.org/osrf/gazebo/pull-requests/2875/example-plugin-that-enables-collisions (accessed Jun. 10, 2019).

[15] "osrf / citysim / plugins — Bitbucket." https://bitbucket.org/osrf/citysim/src/default/plugins/ (accessed Jun. 10, 2019).

[16] URDF, "urdf - ROS Wiki," Jan. 11, 2019. http://wiki.ros.org/urdf (accessed Mar. 29, 2019).

[17] "AutoDrive Challenge - Student Events." https://www.sae.org/attend/student-events/autodrive-challenge/ (accessed Jun. 12, 2019).

[18] "joint_state_publisher - ROS Wiki." http://wiki.ros.org/joint_state_publisher (accessed Jun. 14, 2019).

[19] "Ignition Math: Ignition Math." https://ignitionrobotics.org/api/math/4.0/index.html (accessed Jun. 14, 2019).

[20] T. Foote, "tf: The transform library," in *2013 IEEE Conference on Technologies for Practical Robot Applications (TePRA)*, Apr. 2013, pp. 1–6, doi: 10.1109/TePRA.2013.6556373.

[21] "Gazebo : Tutorial : Gazebo plugins in ROS." http://gazebosim.org/tutorials?tut=ros_gzplugins (accessed Dec. 07, 2019).

[22] "Gazebo : Tutorial : Logical Camera Sensor." http://gazebosim.org/tutorials?tut=logical_camera_sensor&cat=sensors (accessed Jun. 13, 2019).

[23] S. Sharma, G. Tewolde, and J. Kwon, "Behavioral Cloning for Lateral Motion Control of Autonomous Vehicles Using Deep Learning," the 2018 IEEE International Conference on Electro/Information Technology (EIT). IEEE, 2018.

[24] G. A. Kumar, J. H. Lee, J. Hwang, J. Park, S. H. Youn, and S. Kwon, "LiDAR and Camera Fusion Approach for Object Distance Estimation in Self-Driving Vehicles," *Symmetry*, vol. 12, no. 2, Art. no. 2, Feb. 2020, doi: 10.3390/sym12020324.

[25] A. E. AbdELhamed, *AhmedAbdlhamid/Simulation_ENV*. 2019.

Neural Network Based Corn Field Furrow Detection for Autonomous Navigation in Agriculture Vehicles

Niko Anthony Simon
Department of Electrical Engineering
University of St. Thomas
St. Paul MN, United States

Cheol-Hong Min
Department of Electrical Engineering
University of St. Thomas
St. Paul MN, United States

***Abstract*—Row detection in agricultural applications has commonly used Hough transform techniques and traditional signal processing based approaches relating to machine vision. There are various learning based methods available that are capable of producing similar results in terms of detection. In this paper, a neural network based algorithm is developed, and we compare the Hough transform and a machine learning implementation with the proposed approach to determine which would be the most appropriate in a real-time application given a variety of factors including computational performance, accuracy, and environmental variability. Compared to the learning based approaches which rely on training data, Hough transform based detection relies on a variety of processes, including binarization and denoising, which are not required to be explicitly implemented in the machine learning or neural network models. Additionally, to add another layer of diversity to the three possible solutions examined is the consideration for color input data. The Hough transform method and the neural network model implemented both require color input data while the machine learning model relies on texture features instead of color to make its classification predictions. Compared to the traditional image understanding techniques, autonomous vehicles face challenges due to similarities in color and texture between the crops and their surroundings. Therefore, the algorithm is developed to overcome such challenges. Preliminary results show that the neural network model developed was found to offer the most versatility compared to traditional methods and the highest accuracy on the order of 97% for this application across several different input conditions.**

Keywords---Hough Transform, Neural Network, Machine Learning, Machine Vision, Image Processing, Texture Analysis, Color Processing, Autonomous Navigation, Furrow Detection, Agriculture

I. INTRODUCTION

There are a few facets to examine when considering crop row detection using a machine learning model and comparing that with using a neural network model. Traditionally, crop row detection has adopted a purely signal processing approach with a commonly used technique being the Hough Transform. Some considerations when choosing an appropriate solution for a given environment are computation time, reliance on color data, the variability of environment, etc. When deployed in a real-time application and depending on the required resolution, there are several challenges from a design perspective including lighting conditions, complexity and variability of distinct regions, possible similarities between a region of interest versus another region, and unforeseen or periodically occurring phenomena in the input images like a shadow. These aspects may be easier or more complex to address depending on the chosen approach, as suggested by Jinlin in [6].

A major aspect to consider outside of overall computational efficiency is pre-existing knowledge of exactly what the input will contain. For most types of applications, you may have a general idea of what an image input might be, as seen by an autonomous vehicle. However, when operating in a real working environment, there will often be unexpected obstacles or various unforeseen conditions. If everything about the image input is a constant factor, then a machine learning model may be more appropriate given that certain features must be selected to construct the model, which generally suggests some pre-existing knowledge about what exactly will be seen by the system. A neural network, on the other hand, does not require pre-existing information about specific features, rather just an understanding of input dimensionality, resolution and a general expectation of content.

Specifically, in this type of agricultural application, there exists the potential for significant amounts of variation. Identifying areas in the sky region can depend on weather conditions and time of day. This consequentially has an effect on the other regions because if the sky is clear and the sun is out, the grass and crop will be illuminated, but depending on the direction of the sunshine with respect to the crop, it may cast shadows on the grass which looks different from grass in conditions where lighting is not an influence. The crop region is by far the most complex. It is possible when examining close up parts of plant matter belonging to crops to misidentify them as sections of surrounding grass. The anatomy of the corn stalks and the time of the season for harvest also influence the color and texture of the crop, which further complicates classification. The grass region could have patches of dirt, inconsistent grass length, or even unexpected items sitting on the ground. Given the qualities of this type of input data, it might be easy for one to say that a neural network model would be the best pick in every case due to the various complexities however, a machine learning model can be designed with certain specializations in mind that might require a more complex neural network model equivalent. There are too many facets to machine learning and

Figure 1: The three main input samples that will be examined, referred to throughout the analysis as Fig. 1A (left), Fig.1B (center), Fig. 1C (right)

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

neural network models for a logical one to one comparison for this application, therefore this investigation will focus on the specialization that each type of model can offer with an analysis of various techniques and methods seen in [1,3,11,12].

For the machine learning model, feature specialization will be emphasized, and for the neural network approach, model flexibility will be emphasized. This will be shown by examining a machine learning model that used a greyscale input for texture-based identification and a convolutional neural network that uses color input images. To maintain an unbiased comparison between the model performance, the preprocessing and post-processing schemes are kept consistent. The main alternative to these methods is a signal processing approach based on the Hough Transform. Variations of techniques involving the Hough transform have been the most commonly used methods for row detection in agricultural applications. Associated filtering techniques and processes like morphological mathematics are required for this kind of machine-vision based method.

In the comparison of the machine learning and neural network models, three different variations of an input image will be compared. First, Figure 1A is the typical input where the crop is in the expected condition, and there exists some but not overly dramatic distortions in the grass region caused by shadow. The second figure, 1B, is a close-up view of a corn stalk at the ground level. This view contains a minimal amount of the grass region and is highly relevant because the grass and corn regions share several attributes. The program does not have predefined information indicating that it should expect the grass region to be present all the time. If it did, then when facing the crop head-on, it would incorrectly interpret part of the crop region to be the grass region, which would be highly undesirable. Because the crop region is the most varied with several different textures and colors, a close-up view has the potential to further complicate analysis if previous training data was based on the crop observed from a more expected orientation. The third figure, 1C, is an unexpected input on multiple different levels. The lighting and condition of the crop are the main unexpected factors. If the analysis method is insensitive to color, this by itself does not pose a major concern; however, if the crop is in a different condition than was expected based on training data, it may pose identification issues. Additionally, there is an unexpected object on the grass in the distance. In a real-world application, if a farmer were to leave something like farm equipment or an unexpected item in the field, it is important to observe, especially for the training based methods, how the model reacts.

Figure 2: Example division of the input image into 4 equal parts where each part shares the same aspect ratio as the original input

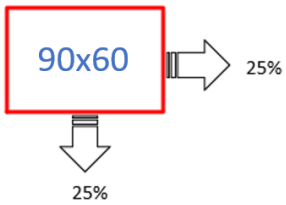


Figure 3: Representation of sliding window that moves across the input image in order to produce segments which feed into the model

II. Pre and Post Processing

There are several aspects to consider in terms of image preprocessing for training based approaches. The raw input image size is 5760x3840, and each input image is further broken up uniformly into segments. This preprocessing method was found to work well during the early stages of development because it is relatively straight forward to use and is capable of setting segment dimensions to their minimum reducible pixel size. As a comparison, preprocessing in the Hough transform based method employs denoising and other filtering techniques and processes the input image as a whole rather than segmenting it piece by piece and assembling a representation of the classification at the output.

An important factor to consider is the size of segments that the input image is divided into and the training images of the same size. Recall that the input image is divided into segments where the model then determines which region of interest they belong to (sky, grass, or corn). When cconsidering possible dimensions for a given segment, the larger the segment, the greater the chance it will include more than one region. However, the smaller the segment, the greater the possibility that it is not distinguishable from other parts from the image. For example, a very small piece of a corn plant image may look indistinguishable from the grass. This means that the segment size is an important design decision, and there are potential tradeoffs to consider, including post-processing.

If segment size is too large using the grid based segmentation method, the resulting predictions will be limited in resolution. Depending on the needs of the autonomous vehicle, a lower resolution might be accurate enough, but if individual sections of the crop at a close up ground level need to be identified, then a more complex preprocessing method is required.

In the case of constructing a training image directory, the segment size must be large enough to contain an identifiable pattern. This means that when using texture to interpret input data, the segments will always need to be larger when compared to using color because a pattern cannot be defined in one pixel.

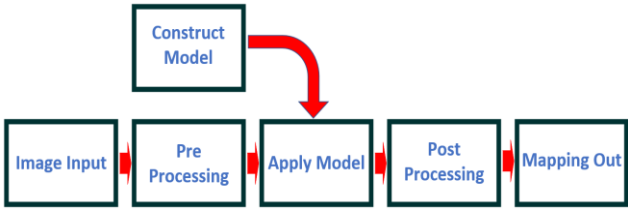


Figure 4: General block diagram for both the machine learning and neural network implementations

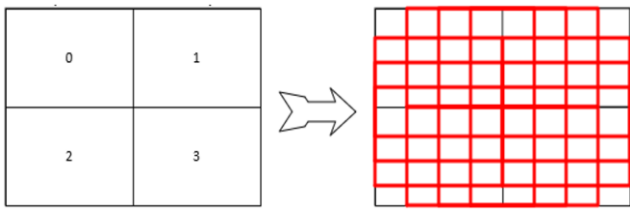


Figure 5: Representation of simple grid division of the input image (left) and generation of overlapping segment of the input image (right)

There is a finite amount of pixels available in each input image. However, it must also be considered that there is a finite amount of area per input image. This is relevant because a pattern takes up much more room than small pixel groupings, which means that later in the program, if the prediction is mapped at a 1:1 ratio using a simple grid division, resolution may be compromised.

Some form of processing needs to happen to the input image to check it one segment at a time with the model. Figure 4 shows the general structure of the implemented training based methods. The model then outputs a prediction of what the correct label is for a given piece. The dimensions of the training images and the input to the model must be the same, and there is a finite number of pieces of the input when using the simple grid division method described previously. Using the second method shown in Figures 3 and 5, the number of pieces the model can process per input is increased, and the accuracy of the overall row detection prediction is increased. This is because any given piece may contain more than one pattern, so a sliding window across would capture more detail in cases where one region begins and another region ends (for example, the boundary between the grass and the corn).

Unlike the Hough based method, a certain amount of post-processing is required due to the fact that the original input is broken up in the preprocessing phase in order to apply the models. The post-processing method is the same for both segmentation schemes discussed in the preprocessing section. The original image input dimensions are related to the number of segments when no overlapping is implemented where:

$$\boldsymbol{H_{original} = H_{segment} * n} \quad (1)$$

$$\boldsymbol{W_{original} = W_{segment} * n} \quad (2)$$

$$\boldsymbol{\#\ of\ segments = n^2} \quad (3)$$

In equations 1-3, $\boldsymbol{n}$ represents the number of segments in each row and column. With overlap, an overlap factor is chosen based on how tightly knit the regions are in the input. If regions are highly intermingled, this factor is increased, which makes it easier for the model to identify regions as a whole because, on a per-segment basis, there is a greater chance of observing a distinct region rather than multiple regions in a given segment.

During development, color filtering techniques were also examined as a possible preprocessing stage. Unlike the Hough transform based method which uses a series of complex filtering techniques in order to identify the crop, a simple mask can be applied to the input based on predefined ranges of color pixel values. If this method worked 100% of the time, it would be an attractive solution to the identification problem, although for this application, two out of the three distinct regions present in the image have overlap in both their color ranges and texture patterns which is demonstrated in Figure 6.

It could be argued that identifying the sky region is not useful if an autonomous vehicle will be traveling on the grass throughout a series of farm fields, however, differentiating the crop from the sky specifically, offers important spatial data which may be relevant for autonomous vehicle orientation and potentially even data which could be used in order to determine the height of a crop from the ground up.

III. Machine Learning Model

The machine learning model consists of a classifier, a pool of training images, and a specific set of features in order to construct a model. For this application, a Random Forrest classifier is used. In testing for this application, the classifier effectiveness of the machine learning model was found to be relatively consistent among commonly used classifier types. The holdout method was applied to verify classifier effectiveness, where data was separated into train and test partitions. Classifier choice is not the main design decision that dictates performance, rather it is the selection of specific features. The machine learning model implements a filter bank of Gabor filters. Gabor filters are a specific type of bandpass filter used to identify various patterns in a given texture. Gabor filters are a type of modulated Gaussian function and have several parameters including wavelength, angle, standard deviation, phase offset, and aspect ratio, shown by Rafi in [5].

The result shown in Figure 7 demonstrates some of the rigidness associated with the need for predefined features as well as a reliance on the training pool. In this case, there is a significant amount of misidentification between the corn and the grass regions and low accuracy in the most critical areas of the image, which are at the boundaries.

Figure 6: Example of color based pixel identification to attempt to separate differentiate regions where the left shows perfect isolation of the sky region and the right shows unsuccessful separation of the crop region due to shared colors and textures with surrounding regions

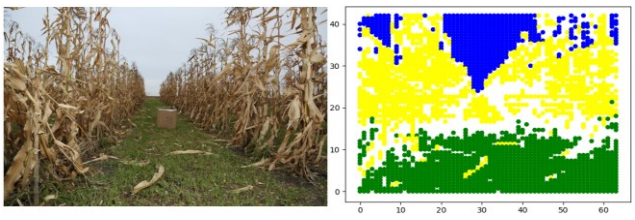


Figure 7: Machine learning model application on input 1C using simple grid division and mapping method set to display predictions in excess of 65% confidence where any classification that falls below this threshold is not mapped, thus creating the white areas which are voids in the mapping

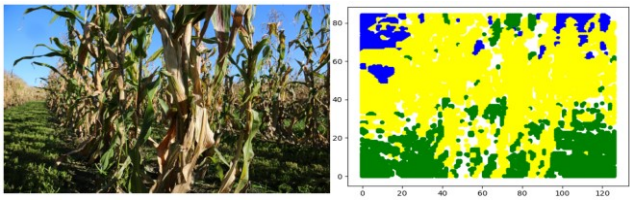


Figure 8: Machine learning model application using overlapping preprocessing method

The result shown in Figure 8 demonstrates better performance with fewer predictions falling short of the set threshold accuracy. However, there is still an undesired area of misclassification of the crop region with a section of the grass region. Even with the implementation of a more complex preprocessing method and considering other small parameters that can be adjusted to tweak accuracy, the main factor influencing performance is still feature diversity. It would be possible to keep adding features and check performance afterward but the associated feature matrix is already significantly large containing several million elements. Adding more features in this way may be inefficient for this application.

Because the input data is two dimensional, the filter parameters also include X and Y coordinates. In order to construct a filter bank, these parameters must be systematically incremented to produce several Gabor filters across the spectrum. It is then possible to identify the filters that contribute most to the feature matrix of the machine learning model.

IV. Neural Network Model

One of the more prominent overarching benefits when comparing using a machine learning model and a neural network model is the reliance on pre-existing knowledge of the system. The machine learning model requires the use of features, and knowing the most appropriate features has an impact on the accuracy. On the other hand, when using the neural network model, certain pre-existing knowledge like features is not required. Instead of features and classifiers being the main construction for the model, parameters and layers are the main focus. First, the three main parameters of interest are the training cycle (epochs), iterations, and batch size. When changing these parameters, the size of the model is a consideration. Similar to the machine learning model, the neural network model is saved after it is trained to prevent having to train data multiple times during normal operation. In the case of the machine learning model, the size of the saved model is proportional to the number of features in the feature matrix, whereas, in the neural network model, the memory used can be effected by the number of training cycles. Similarly, the batch size dictates how many sets of samples will be processed at a time for training. Smaller batch size will train faster and require less memory where a larger batch size will ideally have greater accuracy at the cost of training speed and memory resources.

Getting a clear sense of the effect of these parameters is important, so there is a benefit to examining data related to varying these parameters and the impact on the model. Changing the layers and changing these three main parameters are the main ways to influence the effectiveness of the model. There are several aspects to consider when designing the model, but the main ones are the activation function, optimizer, and loss function. Secondary factors to consider are training cycles and batch size. A discussion of the relevant layer types used for this model is important to understand its overall effect. A dropout layer is meant to reduce the chance of developing a model that overfits to training data by randomly removing nodes. This, in effect, regularizes the network by decreasing its size. It means that any single feature cannot be relied upon because it could be randomly eliminated from the model. An unbalanced scheme that is overly dependent on one subset of features would be detected by observing the effects produced by applying this layer to a model. While the dropout layer works to prevent overfitting, pooling layer downsamples, as shown by Kim [10].

The next relevant neural network functionality that needs to be addressed is pooling. When applied to an image, pooling downsamples the input, by reducing the number of pixels systematically. Effectively the pooling operation works like a sliding window across the input. By reducing the resolution of the data, this helps to decrease the number of weights required.

The convolution layer detects the desired patterns in images where a filter and convolution window transform the input. This is essentially just scanning over image pixels of predefined window size and calculating the dot product per each window. The loss function and optimizer work together to minimize loss through the best possible calculation of model weights, while the purpose of an activation function is essentially just a transform from an input to the output of a neuron.

The RMSprop optimizer is essentially the moving average of the square of gradients. Comparable to Adam, which is regarded as the most used optimizer, RMSprop is one of the most widely used. The choice of the optimizer is highly application-specific, and through testing, RMSprop was found to deliver consistent results. This is most likely due to the averaging operation applied by RMSprop rather than the gradient descent, which is used by Adam.

Sparse categorical cross-entropy is used for the loss function of the neural network model. Categorical cross-entropy based loss functions are appropriate because they are designed for single label categorization. In an application where a particular region could be categorized as more than one thing, this type of loss function would not be appropriate. For example, it would not be logical for the grass to only sometimes be categorized as grass. The Softmax layer increases maximum values and decreases minimum values, in effect normalizing the values so that probabilistic functions can be applied to the data.

Relu, (Rectified Linear Unit) is one of the most commonly used activation functions in neural networks, mainly for its general performance in both quality and computational performance. Other activation functions might produce slightly better performance for this application at the tradeoff of being more computationally intensive. In a real-time application, identifying all possible sources of preventable lag are critically important. Relu passes all inputs above zero and suppresses all inputs below zero. This design choice would also become more relevant if a deep neural network were to be considered instead. Compared to activation functions like a hyperbolic tangent, which are fundamentally bounded mathematically, Relu performs better in cases where many layers introduced gradient related issues for the performance of a model shown by Guo [9].

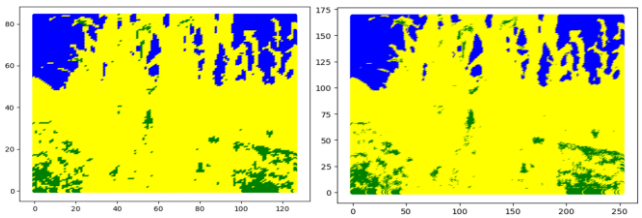


Figure 9: Overlap factor of 2 (left) and overlap factor of 4 (right) for 1B showing enhanced resolution and clear separation of crop and sky

V. RESULTS

When comparing the training based (machine learning and neural network) approaches specifically, it is important to compare the preprocessing methods which offer varying levels of output resolution. As an example, Figure 9 shows the effect of increased overlapping in the preprocessing stage on the same input. The result is a more defined resolution at the cost of computation time, and the same results are observed when using the machine learning model. The main effect of the preprocessing method for both training based model types comes down to resolution. Minimum resolution with consistent accuracy may be a suitable solution for this type of application. Over-processing the input to achieve higher resolution sacrifices real-time performance for possibly unneeded accuracy.

Regarding the state of prediction output for each model, how neural network model output is relatively straight forward as far as expected accuracy while the machine learning model output requires some interpretation of the predictions before a final output is established. Training pool size and the reliance on training data is also a consideration. A diverse and extensive pool of training data is required for the machine learning and neural network models, while the Hough transform based method does not require this.

Figure 10 shows some of the challenges when designing the neural network model for this application. On the left is an application of the neural network model with relatively high accuracy on the order of 97% while on the right is the same model with the use of a different optimizer. This creates near-perfect classification in two of the three regions with mass misclassification in the other region. Model reusability is also a consideration compared to Hough transform based methods, which may require the change of several parameters with a change in environment. If expected crop conditions are known beforehand, different models can be swapped out as needed for changes in the seasons.

Another factor to consider is the actual application in the field in terms of real-time response. Both the machine learning

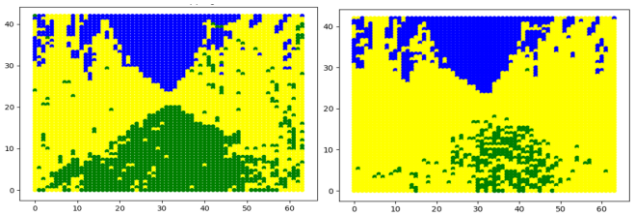


Figure 10: Example of neural network model tuning for input Fig. 1A The result on the left is relatively consistent and the result on the right exhibits a large amount of misclassification but only in one region

TABLE I. TABLE OF OPTIMIZER EFFECTIVENESS

SGD	***Not Effective***
Adadelta	***Somewhat Effective***
Adagrad	***Somewhat Effective***
Adam	***Somewhat Effective***
Adamax	***Somewhat Effective***
Nadam	***Not Effective***
RMSprop	***Highly Effective***

and neural network applications are GPU intensive. Also, the Hough transform application relies on several predefined signal processing functions. All of these have the potential to be computationally intensive when running with onboard hardware and not in a simulation.

Another similar application explained by Cheng in [2] where we can draw some parallels is detecting curbs on roads and sidewalks for autonomous vehicles. Curb detection could employ classifier based techniques, along with geometry-based techniques. Unlike for curb detection, this application does not have regular distinct and reliable geometries to rely on, rather it contains loosely expected textures and colors. The Hough transform based method relies on the input image's pixel color data converted to the HSV (Hue, Saturation, Value) color space after noise removal and binarization operations, the data is in a state where the Hough transform can be applied. The Hough transform essentially maps points in the image to a parameter space to construct decision boundaries. Depending on application specifics, there could be a few issues with using this method, as suggested by Tu in [7] and Zheng in [8].

It could be argued that using a Hough transform based method could be performed more efficiently in terms of computational intensiveness because of the difference in preprocessing compared to the training based methods. In the example of curb detection previously mentioned, identifying the vanishing point in the input image is an important part of the program function in order to establish a sense of depth and forward target. For the neural network and machine learning implementations, in an expected orientation, the vanishing point could be implicitly determined by the point at which all three regions meet.

Shown in Figure 11 is the ideal result for the binarization stage, where the crop is clearly distinguished. In cases where the positioning of the crop or the orientation of the camera is more complex, binarization results become less defined, which results in the calculated decision bounds becoming less reliable.

In cases where the autonomous vehicle might make a turn, the camera input may be in an orientation where it is head-on

Figure 11: Hough transform preprocessing binarization showing crop and grass separation with the sky seen as the same type of area as the grass

with one of the regions rather than observing it from an expected viewpoint. When analyzing this type of input, the Hough transform becomes more inconsistent in the binarization stage. As shown in Figure 12, the grass and corn regions are heavily intermingled, which ultimately makes it difficult to form a decision boundary. A neural network scheme can more accurately differentiate regions without any other functionality or filtering stages required.

When analyzing the three distinct variations of input images, it becomes clearer what the possible strengths and drawbacks are for each method. In the case of inputs 1C and 1A, which exhibit relative uniformity in orientation, the Hough transform based method is highly applicable. Input 1C poses the same issues with the training based methods due to the unexpected change of one of the regions.

The binarization stage of the Hough transform based detection method is consistent for uniform rows at a uniform perspective. However, any number of factors having to do with changing the input may introduce undesired effects. The major strength of the Hough transform based method is that it doesn't require a pool of training data or extensive post-processing. The drawback is in its lack of flexibility under certain input conditions where it performs very well in an expected orientation but would require extra functionality added if certain environmental factors changed from one input image to the next.

Regarding the neural network scheme, as long as it has a robust training pool, it is highly adaptable in catching unexpected input conditions. In contrast, the machine learning model, whose inherent issue is establishing the proper features, still maintains this adaptive quality with a robust training pool. Also, the HSV color space is somewhat resilient in differentiating lighting conditions as seen in [4] by Kaur, however, with changing grass or crop color throughout the season or even operation during dusk or evening times, the flexibility of multiple different models that can be swapped out as needed becomes an option worth considering.

VI. CONCLUSIONS AND FURTHER WORK

The paper presents preliminary analysis and results of a Neural Networked based algorithm to detect rows so that farm vehicles can autonomously navigate through corn fields without damaging the crops. Challenges due to color and texture similarities in the crop rows were addressed and successfully resolved in the preliminary analysis. In order to rival the performance seen by the neural network model, the machine learning scheme requires several optimizations, including the consideration of additional features. Also, the Hough transform based identification isolates the crop region by differentiating it from everything else, but unlike the machine learning and neural network approaches, it does not offer specific information on the other surroundings like the sky. Even though the Gabor filter is an excellent feature to identify textures, other factors like optimal classifier choice, the calculation time for higher resolution data, and overall resilience of model performance for a given input clearly suggest that for this type of complex classification problem, a neural network model is easier to implement. Further work will consist of modifying the preprocessing stage to allow for video input and increasing the complexity of the neural network implementation. The algorithm will be further tested on other crops such as soybeans to test the stability of the algorithm.

Figure 12: Hough transform preprocessing binarization of input example 1B showing indistinct separation results

REFERENCES

[1] Y. Li, J. Zhang, P. Gao, L. Jiang and M. Chen, "Grab Cut Image Segmentation Based on Image Region," *2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC)*, Chongqing, 2018, pp. 311-315, doi: 10.1109/ICIVC.2018.8492818.

[2] M. Cheng, Y. Zhang, Y. Su, J. M. Alvarez and H. Kong, "Curb Detection for Road and Sidewalk Detection," in *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10330-10342, Nov. 2018, doi: 10.1109/TVT.2018.2865836.

[3] M. Chanda and M. Biswas, "Plant disease identification and classification using Back-Propagation Neural Network with Particle Swarm Optimization," *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, 2019, pp. 1029-1036, doi: 10.1109/ICOEI.2019.8862552.

[4] M. Kaur and C. Min, "Automatic Crop Furrow Detection for Precision Agriculture," *2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS)*, Windsor, ON, Canada, 2018, pp. 520-523, doi: 10.1109/MWSCAS.2018.8623906.

[5] M. Rafi and S. Mukhopadhyay, "Segmentation of Synthetic Textures Employing Gabor Filter Magnitude in a Multi-Channeling Environment," *2017 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, Jaipur, 2017, pp. 269-274, doi: 10.1109/SITIS.2017.53.

[6] X. Jinlin and J. Weiping, "Vision-Based Guidance Line Detection in Row Crop Fields," *2010 International Conference on Intelligent Computation Technology and Automation*, Changsha, 2010, pp. 1140-1143, doi: 10.1109/ICICTA.2010.400.

[7] C. Tu, B. J. van Wyk, K. Djouani, Y. Hamam and S. Du, "An efficient crop row detection method for agriculture robots," *2014 7th International Congress on Image and Signal Processing*, Dalian, 2014, pp. 655-659, doi: 10.1109/CISP.2014.7003860.

[8] L. Zheng and J. Xu, "Multi-crop-row detection based on strip analysis," *2014 International Conference on Machine Learning and Cybernetics*, Lanzhou, 2014, pp. 611-614, doi: 10.1109/ICMLC.2014.7009678.

[9] T. Guo, J. Dong, H. Li and Y. Gao, "Simple convolutional neural network on image classification," *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*, Beijing, 2017, pp. 721-724, doi: 10.1109/ICBDA.2017.8078730.

[10] B. Ko, H. Kim and H. Choi, "Controlled dropout: A different dropout for improving training speed on deep neural network," *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Banff, AB, 2017, pp. 972-977, doi: 10.1109/SMC.2017.8122736.

[11] F. Zou, L. Shen, Z. Jie, W. Zhang and W. Liu, "A Sufficient Condition for Convergences of Adam and RMSProp," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 2019, pp. 11119-11127, doi: 10.1109/CVPR.2019.01138.

[12] M. C. A. Venal, A. C. Fajardo and A. A. Hernandez, "Plant Stress Classification for Smart Agriculture utilizing Convolutional Neural Network - Support Vector Machine," *2019 International Conference on ICT for Smart Society (ICISS)*, Bandung, Indonesia, 2019, pp. 1-5, doi: 10.1109/ICISS48059.2019.8969799.

TiO_2-GO Field Effect Transistors for Amplified Ethanol Sensing

Teena Gakhar and Arnab Hazra*

Dept. of Electrical & Electronics Engineering, Birla Institute of Technology and Science (BITS)-Pilani, Vidya Vihar, Rajasthan 333031, India

*Email: arnabhazra2013@gmail.com , arnab.hazra@pilani.bits-pilani.ac.in,
Tel: +91-1596-255724,

Abstract- In this work we have proposed *p*-TiO_2-GO nanocomposite field effect transistor based ethanol sensor. *p*-type TiO_2 nanoparticles was prepared by sol-gel method and mixed with 2 wt% aquas solution of graphene oxide (GO) and sonicated for 30 min. The nanocomposite was prepared in combination of 5 vol% *p*-TiO_2 nanoparticles with 95 vol% GO. The morphological and structural characterizations of developed nancomposite were carried out with field emission scanning electron microscopy (FESEM) and Raman spectroscopy techniques, respectively. The *p*-TiO_2-GO field effect transistor (FET) sensor showed a response magnitude of 6% in terminal structure when V_{GS}=0 and 41% as in three terminal structure when V_{GS}=0.65 V in the exposure of 100 ppm ethanol at 100°C. The *p*-TiO_2-GO FET showed maximum ~7 times amplification in sensitivity due to application of positive gate voltage.

Keywords- TiO_2-GO hybrid, FET, Ethanol sensing, Amplified sensitivity

I. INTRODUCTION

Highly stable and low power gas/vapor sensors along with low operating temperature is of great interests for the researchers. A number of experiments and techniques have been reported to develop a highly efficient vapor sensor [1]. Ethanol is one of the important volatile organic compounds (VOCs) used in multiple industrial and domestic applications. Low ppm and low temperature detection of ethanol is important to monitor the quality of alcoholic beverages, restrict drunk driving, diagenesis diseases though breath analysis [2].

Different materials and methods have already been implemented for efficient detection of ethanol by using solid state vapor sensors. Graphene and its derivatives like graphene oxide, reduced graphene oxides are potential candidates are reported for gas/vapor sensing in recent years. Graphene oxide (GO) in composites with metal oxides nanoforms attract a significant attention in recent days. Not only the gas/vapor sensing applications, GO-metal oxide nanocomposites have proven themselves as a promising candidate in different applications like photocatalysis [3], solar cells [4, 5] etc. The nanocomposites have many advantages like high mechanical strength, high specific area, low activation and sufficient fictional groups. Also, graphene serves as strong substrate in the formation of nanocomposite [6]. The formation of p-n junction on the synthesis of hybrid leads to a better performance than pure material [7]. The hybrid provides excellent chemical and electrical properties which in turn improves the sensing [8]. Various reports have been published stating the improvement in sensor performance by united effect of both graphene oxide and metal oxide. Rare experiments have been done by the researchers in the field of graphene field effect transistor for gas/VOC sensing. Nakaruma group have reported graphene FET based gas sensor with Si/SiO_2 surface having enhanced surface adsorption sites [9]. Luszczek and co-workers reported the modelling of a top gated graphene field effect transistor (GFET) which was suitable for electronic sensing application [10]. Liu and group synthesized single graphene field effect transistor (FET) which showed selective sensing towards NO_2 NH_3, H_2O and CH_3OH [11]. Lu and group have demonstrated the ultra-fast ammonia sensing by fabricating positively gated reduced graphene oxide field effect transistor. They have reported fast, repeatable and room temperature gas sensing [12]. Rumyantsev and co-workers have showed selective sensing of individual gases by using graphene devices. They demonstrated the sensing in terms of two parameters - low frequency noise and memory step [13]. Inaba and co-workers have fabricated graphene field effect transistor gated by ionic liquid for ammonia sensing [14]. Larisika and group have reported olfactory biosensor based on reduced graphene oxide field effect transistor [15].

Until now, the nanocomposites of graphene, graphene oxide or reduced graphene oxide with metal oxides in FET structure for gas/VOC sensing has not have been explored by the researchers. Herein, we have investigated the FET behaviour of nanocomposite for the efficient detection of ethanol. *p*-TiO_2 nanoparticles

978-1-7281-9615-2/20/$31.00©2020 IEEE

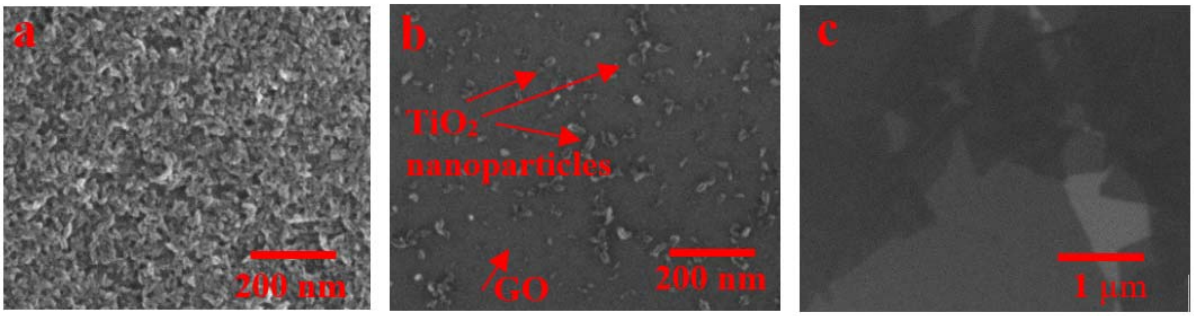


Fig. 1 FESEM images of (a) pure *p*-TiO_2 nanoparticles (S_1), (b) *p*-TiO_2 nanoparticles-GO hybrid (S_2), (c) GO nanosheet (S_3).

decorated graphene oxide field effect transistor were fabricated for ethanol sensing. An amplified sensor response was achieved with an accurate gate voltage (V_{GS} = 0.65 V) which was 7 times more than the sensor response obtained in the two-terminal planar configuration.

I. EXPERIMENT

p-TiO_2 nanoparticles were fabricated by sol-gel method where 2 ml of titanium iso-propoxide (TTIP) was added into acetic acid dropwise. This solution was kept inside an ice bath under stirring condition. After that 40 ml of deionized water was added to the mixture of TTIP and acetic acid and was kept under continuous stirring for 90 min. As-prepared TiO_2 solution was kept in an unilluminated place for aging which lead to the change in the colour of the TiO_2 solution from translucent to hazy [16]. Dense (6.2g/L) graphene oxide solution, purchased from graphene supermarket was used. 0.2 wt% aqueous solution of graphene oxide was prepared for the formation of nanocomposite. To synthesize the nanocomposite, 5 vol% of GO solution and 95 vol% of pure *p*-TiO_2 nanoparticles solution were mixed properly through sonication.

FET sensors were fabricated with pure *p*-TiO_2 nanoparticles (S_1), *p*-TiO_2-GO nanocomposite (S_2) and pure GO solution (S_3). All the three solutions were deposited on SiO_2/Si (500 µm thick <100>) substrate with the help of dip coating technique (Appex Instruments: Xdip-SV-1). This SiO_2/Si wafer which was taken as the substrate was cleaned with ethanol and was cut into 1cm × 1 cm for sample formation. Gate contact was taken from the back side of the substrate after selective HF etching of SiO_2 layer. All the three samples were kept for annealing at 250 □C for 5 hours.

Morphological characterization of all the three samples (pure *p*-TiO_2 nanoparticles, pure GO solution and the *p*-TiO_2-GO nanocomposite) were done through FESEM (Fig.1). Structural characterization was confirmed by Raman spectroscopy (Fig.2).

All the fabricated sensors were tested for ethanol sensing in a sealed glass chamber (450 ml) with a gas inlet, outlet and electrical measurements facilities. The sensor setup has already been explained in our previous reports [17],[18]. Dry synthetic air (N_2:O_2=4:1) was used as the carrier gas and a constant flow of 450 sccm was used in throughout sensor study. The response magnitude of each sensor was calculated using $[(I_a-I_g)/I_a]x100$ where I_a is the current of the sensor measured in air ambient and I_g is current of the sensor measured in ambient.

III. RESULTS AND DISCUSSIONS

(a) Material Characterization

A dense and continuous layer of pure *p*-TiO_2 nanoparticles was observed in S_1 sample by FESEM (Fig1.a). The size of the TiO_2 nanoparticles synthesized was approximately 10 nm. The uniform decoration of *p*-TiO_2 nanoparticles on GO nanosheets is seen the nanocomposite formed (S_2) (Fig.1b). Scanning electron micrograph also confirmed that the nanocomposite formed has not affected the morphology and structure of both *p*-TiO_2 nanoparticles and GO (Fig.1b).

Structural characterization of the three samples were studied using Raman spectroscopy which is depicted through two scans; one from 100 cm^{-1} to 700 cm^{-1} (for anatase TiO_2) and the other from 1200 cm^{-1} to 1800 cm^{-1} (for GO). Raman spectroscopy was done using the green laser (500 nm) at room temperature (300 K). The anatase crystallinity of *p*-TiO_2 nanoparticles in S_1 is verified by the modes E_g (144 cm^{-1}), E_g (199 cm^{-1}), B_g (399 cm^{-1}) and E_g (639 cm^{-1}) (Fig2.a) [19]. The sharp intensity peaks D and G in pure graphene oxide sample (S_3) are observed very clearly. The *p*-TiO_2 nanoparticles-GO hybrid sample (S_2) comprises of very low intensity E_g (144 cm^{-1}), D and G peak which confirms the presence of both in the TiO_2 and GO. A sharp peak around 525 cm^{-1} is appearing in all the three samples due to the silicon of SiO_2/Si substrate.

(b) Sensor study in two terminal structure (V_{GS}=0)

All the fabricated sensors i.e. pure *p*-TiO_2 nanoparticles, pure GO and the *p*-TiO_2-GO nanocomposite FETs were tested for ethanol

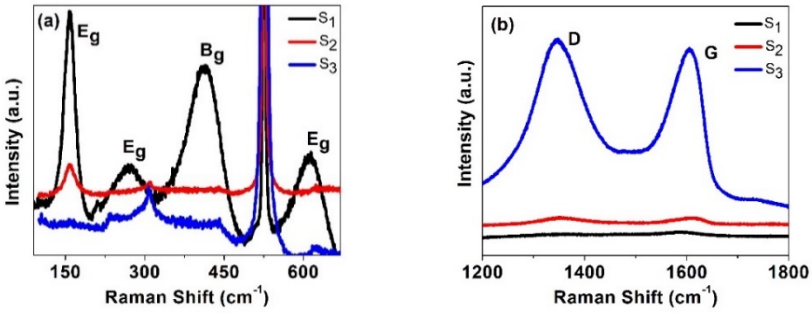


Fig. 2. Raman spectroscopy study in the wave number of (a) 100-700 cm^{-1}, (b) 1200-1800 cm^{-1}.

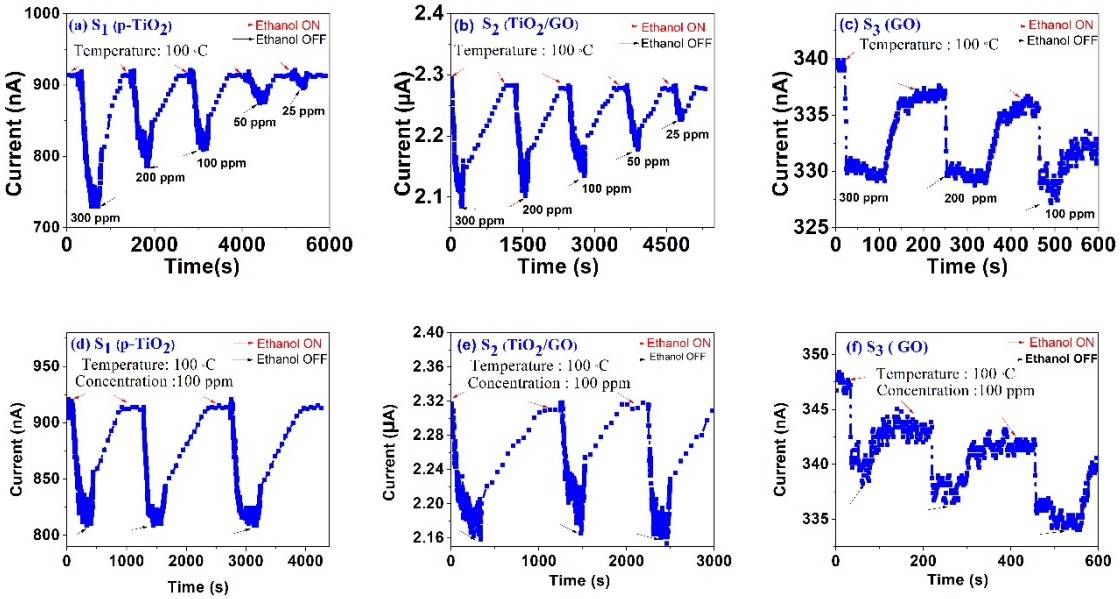


Fig.3 Transient behaviour in two terminal structure of the sensors (V_{GS}=0 V) for (a) *p*-TiO_2 nanoparticles, (b) *p*-TiO_2 nanoparticles-GO nanocomposite and (c) GO, Repeated cycles study in the exposure of 100 ppm ethanol for (d) S_1 (e) S_2 and (f) S_3

sensing at 100 □C. The base line resistance of *p*-TiO_2 nanoparticles and pure GO was 550 KΩ and 1.44 MΩ respectively at 100 □C. The *p*-TiO_2-GO nanocomposite showed a resistance of 200 KΩ. This drop in the base line resistance of the synthesized nanocomposite is due to the formation of Ti-O-C bond and the transfer of holes from *p*-TiO_2 nanoparticles to GO upon formation of nanocomposite [20]. The nanohybrid acted like C doped TiO_2 and thus exhibited a higher stability and higher conductivity [21]. The transient response of all the three samples were taken in two terminal structure i.e. V_{GS}=0 in the range from 300 ppm to 25 ppm and shown in Fig. 3. *p*-TiO_2 sensor (S_1) exhibited relatively high response (~30% in 300 ppm) as compared to others (Fig. 3a). The response magnitude of pure GO sample (S_3) was very low among all the samples and also this sample was not able to sense below 100 ppm (Fig. 3c). S_2 had shown a moderate response magnitude of ~8% in 300 ppm ethanol at V_{GS}=0 (Fig. 3b). The *p*-type behaviour of pure TiO_2 nanoparticles were due to the presence of excess oxygen and deficiency of vacancy (Vo) type defects. Holes as the majority carrier was confirmed by the reduction of current on exposure to reducing type ethanol vapours. [22]. Also, adsorbed oxygen on the surface of graphene oxide extinct the electrons in GO which is responsible for its *p*-type behaviour [23],[24]. The *p*-TiO_2-GO hybrid also showed the *p*-type behaviour during ethanol sensing. On exposure to a reducing vapour i.e. ethanol, electrons are attracted towards the surface of *p*-TiO_2 nanoparticles and GO resulting in a recombination process of holes and electrons and thus current is decreased which is depicted in transient (Fig. 3). The stability behaviour i.e. repeated cycle study all the three sensors i.e. *p*-TiO_2 (S_1), *p*-TiO_2/GO nanocomposite (S_2) and pure graphene oxide (S_3) are shown in Fig.3.d-f. The reapted cycles for pure graphene oxide sample depicted a lower drift idicating less stability behaviour as compared to the *p*-TiO_2/GO nanocomposite.

(c) Amplified sensitivity in FET structure (V_{GS}>0)

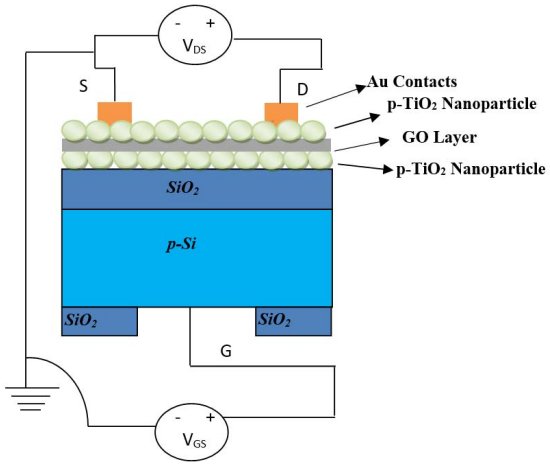


Fig.4 Schamatic of *p*-TiO_2/GO FET with essential biasing arrangement.

The three terminal device structure with applied back gate potential (V_{GS}) is shown in Fig.4. All the above transients were taken in the two- terminal configuration with no gate voltage applied (V_{GS}= 0). Being a pure metal oxide, S_1 was unable to show any FET behaviour i.e. V_{GS} effect. The devices S_2 and S_3 in three terminal configurations showed a V shaped current – voltage (I_{DS}-V_{GS}) characteristics with a minimum drain current point known as the Dirac point (V_{Driac}). I_{DS}-V_{GS} characteristics were taken at different drain voltages (V_{DS} = 0.2 V – 1 V) to analyse the ambipolar nature of graphene oxide and structure (Fig.5).

The majority carrier concentration *p*-TiO_2-GO nanohybrid in field effect devices can be altered from holes to electrons in the 2D channel with the help of applied suitable gate potential. S_1 the pure TiO_2 nanoparticles sample did not show any ambipolar behaviour and S_3 showed very low unstable response (Fig. 3c,f). So, the *p*- TiO_2-GO nanocomposite sample was taken for sensor study only.

I_{DS}-V_{GS} characteristics in air ambient and 100 ppm ethanol for the nanocomposite was measured at V_{DS}= 0.5 V and at 100 □C (Fig6.a) and simultaneously the response magnitude was calculated for different V_{GS} (Fig.6.b).The response magnitude was calculated using the formula [($I_{DS\ air}$ - $I_{DS\ ethanol}$)/$I_{DS\ ethanol}$]×100. At V_{GS}= 0.65 V response spike of 60% magnitude was obtained for the nanocomposite sample. After that the transient was measured at V_{GS} = 0.65 V and V_{DS}=0.5 V from the range of 300 ppm to 25 ppm (Fig6.c). The sensitivity obtained was 41 % in exposure to 100 ppm ethanol at 100 □C which was 7 times higher than the response magnitude obtained with two terminal devices.

IV. CONCLUSION

Herein, we have prepared three sensors were prepared i.e. pure *p*-TiO_2 nanoparticles, pure GO and *p*-TiO_2-GO nanocomposite. *p*-TiO_2 nanoparticles- GO field effect transistor exhibited extraordinary high sensitivity at appropriate gate voltage. FESEM confirmed the decoration of *p*-TiO_2 nanoparticles on graphene oxide without any structural deformation. Raman spectroscopy showed the anatase crystallinity of TiO_2 nanoparticles and presence of graphene oxide in the nanocomposite. The ambipolar nature of graphene oxide is depicted through the I_{DS}-V_{GS} characteristic. In conclusion, an amplified sensor response has been achieved by with *p*-TiO_2-GO field effect structure by applying a specific V_{GS} measured from the two I_{DS}-V_{DS} characteristics in air and VOC ambient respectively. An amplified transient response measured by S_2 at V_{GS} = 0.65 V was 7 times more than the transient obtained in two terminal configuration (V_{GS} = 0 V) from 300 ppm to 25 ppm at 100 □C.

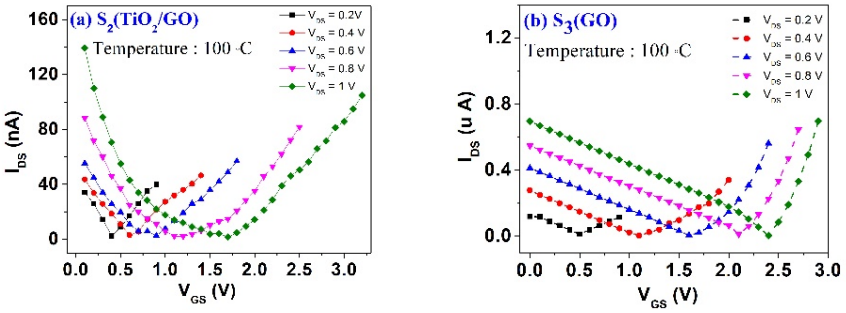


Fig. 5. I_{DS}-V_{GS} characteristics varying V_{DS} from 0.2 V to 1 V. (a) *p*-TiO_2-GO nanocomposite FET, (b) pure GO-FET

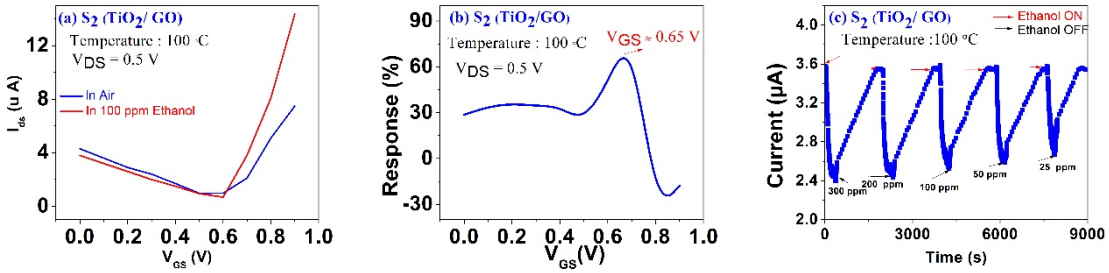


Fig.6. *p*-TiO_2-GO FET. (a) I_{DS}-V_{GS} characteristics in air and ethanol, (b) measured response magnitude from I_{DS}-V_{GS} characteristics, (c) amplified transient response S_2 at V_{GS}=0.65 V.

References

1. X. Li, Y. Zhao, X. Wang, J. Wang, A. M. Gaskov, and S. A. Akbar, "Reduced graphene oxide (rGO) decorated TiO_2 microspheres for selective room-temperature gas sensors," Sensors and Actuators B: Chemical, vol 230, pp.330-336, 2016.
2. J. Liu, X. Wang, Q. Peng, Y. Li, "Vanadium Pentaoxide Nanobelts: Highly selective and stable ethanol sensor material," Advanced Materials, vol. 17 no.6, pp. 764 – 767, 2005.
3. S. Vaclav, D. Popelková, and Petr Vlácil, "TiO_2–graphene nanocomposite as high performace photocatalysts," The Journal of Physical Chemistry C, vol115, no. 51, pp. 25209-25218, 2011.
4. S. P. Lim, A. Pandikumar, N. M. Huang, and H. N. Lim, "Reduced graphene oxide–titania nanocomposite□modified photoanode for efficient dye□sensitized solar cells," International Journal of Energy Research, vol 39, no. 6, pp. 812-824, 2015.
5. Z. Ye, H. Tai, R. Guo, Z. Yuan, C. Liu, Y. Su, Z. Chen, and Y. Jiang, "Excellent ammonia sensing performance of gas sensor based on graphene/titanium dioxide hybrid with improved morphology, " Applied Surface Science, vol 419, pp. 84-90, 2017.
6. Varghese, Seba S., Sunil Lonkar, K. K. Singh, Sundaram Swaminathan, and Ahmed Abdala, "Recent advances in graphene-based gas sensors," *Sensors and Actuators B: Chemical* 218 pp. 160-183, 2015.
7. S. G. Chatterjee, Somenath Chatterjee, Ajoy K. Ray, and Amit K. Chakraborty, "Graphene–metal oxide nanohybrids for toxic gas sensor: a review," *Sensors and Actuators B: Chemical* 221, 1170-1181, 2015.
8. D. Guo, P. Cai, J. Sun, W. He, X. Wu, T. Zhang, X. Wang, and X. Zhang, "Reduced-graphene-oxide/metal-oxide pn heterojunction aerogels as efficient 3D sensing frameworks for phenol detection," Carbon, vol 99 pp. 571-578, 2016.
9. M. Nakamura, Y. Kanai, Y. Ohno, K. Maehashi, K. Inoue, K. Matsumoto, "Graphene-FET-based gas sensor properties depending on substrate surface conditions," Japanese Journal of Applied Physics, vol 54, pp. 06FF11, 2015.
10. M. Łuszczek, M, Turzyński, D. Świsulski, "Modelling of Graphene Field-Effect Transistor for electronic sensing applications," Przegląd Elektrotechniczny, vol 91, pp.170-172, 2015.
11. Y. Liu, S. Lin, L. Lin, "A versatile gas sensor with selectivity using a single graphene transistor," In 2015 Transducers-2015 18th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS) IEEE, pp. 961-964, 2015.
12. G. Lu, Kehan Yu, Leonidas E. Ocola, and Junhong Chen, "Ultrafast room temperature NH_3 sensing with positively gated reduced graphene oxide field-effect transistors," *Chemical Communications* 47, no. 27 pp. 7761-7763, 2011.
13. S. Rumyantsev, Guanxiong Liu, Radislav A. Potyrailo, Alexandr A. Balandin, and Michael S. Shur, "Selective sensing of individual gases using graphene devices," *IEEE Sensors Journal* 13, no. 8, pp. 2818-2822, 2013.
14. A. Inaba, Kwanghyun Yoo, Yusuke Takei, Kiyoshi Matsumoto, and Isao Shimoyama, "Ammonia gas sensing using a graphene field–effect transistor gated by ionic liquid," *Sensors and Actuators B: Chemical* 195 pp. 15-21, 2014.
15. M. Larisika, C. Kotlowski, C. Steininger, R. Mastrogiacomo, P. Pelosi, S. Schütz, S. F. Peteu, C. Kleber, C.R. Rozman, C. Nowak and W. Knoll, "Electronic olfactory sensor based on A. mellifera odorant□binding protein 14 on a reduced graphene oxide field□effect transistor," *Angewandte Chemie International Edition* 54, no. 45 pp. 13245-13248, 2015.
16. A. Hazra, P. Chattopadhyay, and P. Bhattacharyya, "Hybrid Fabrication of Highly Rectifying $ p $-$ n $ Homojunction Based on Nanostructured TiO_2," IEEE Electron Device Letters, vol 36, no. 5 pp. 505-507, 2015.
17. P. Bindra, A. Hazra, "Selective detection of organic vapors using TiO_2 nanotubes based single sensor at room temperature," Sensors and Actuators B: Chemical, vol 290, pp. 684-690, 2019.
18. T. Gakhar, A. Hazra, "Oxygen vacancy modulation of titania nanotubes by cathodic polarization and chemical reduction routes for efficient detection of volatile organic compounds," Nanoscale vol 12, pp. 9082-9093, 2020.
19. P. Wang, J. Wang, X. Wang, H. Yu, J. Yu, M. Lei, Y. Wang, "One-step synthesis of easy-recycling TiO_2-rGO nanocomposite photocatalysts with enhanced photocatalytic activity," Applied Catalysis B: Environmental vol 132 pp. 452-459, 2013.
20. H. Zhang, X. Wang, N. Li, J. Xia, Q. Meng, J. Ding, J. Lu, "Synthesis and characterization of TiO_2/graphene oxide nanocomposites for photoreduction of heavy metal ions in reverse osmosis concentrate," RSC advances vol 8, pp. 34241-34251, 2018.
21. Y. Zhang, C. Pan, "TiO_2/ Graphene composite from thermal reaction of graphene oxide and its photocatalytic activity in visible light," J. Matter Science, vol. 46, pp. 2622-2626, 2011.
22. A. Hazra, S. Das, J. Kanungo, C. K. Sarkar, S. Basu, "Studies on a resistive gas sensor based on sol–gel grown nanocrystalline p-TiO_2 thin film for fast hydrogen detection," Sensors and Actuators B: Chemical, vol 183, pp. 87-95, 2013.
23. O. Okobiah, R. F. Reidy, Surface interactions: functionalization of graphene oxide and wetting of graphene oxide and graphene," Current Organic Chemistry, vol. 20, pp. 674-681, 2016.
24. S. Perumbilavil, P. Sankar, T. P. Rose, R. Philip, "White light Z-scan measurements of ultrafast optical nonlinearity in reduced graphene oxide nanosheets in the 400–700 nm region," Applied Physics Letters, vol.107, pp. 051104, 2015.

Method for Reducing Load and Generator Output In Case of Transmission Line Overload

Chokder Alamgir Bartend Russell
Dept. of Electrical & Electronic Engg.
Independent University, Bangladesh
Dhaka, Bangladesh
cabr1234@gmail.com

Shahriar Khan
Dept. of Electrical & Electronic Engg.
Independent University, Bangladesh
Dhaka, Bangladesh
skhan@iub.edu.bd

***Abstract*—In spite of advances in electrical and computer engineering, the issue of blackouts in the power grid has not been solved. A new method is proposed to alleviate transmission line overloads, created by the outage of a single transmission line. An overloaded transmission line may trip, causing power to be redistributed to other transmission lines, which may get overloaded themselves. This may lead to cascading outages, destabilization, and blackouts. The proposed new method uses the PSS®E to shed load from specific buses and to reduces power output of specific generators. The method was applied to the IEEE 39-bus power system using additional softwares Python, and Excel. All 34 transmission lines were disconnected individually and transmission lines were monitored for overcurrents. The proposed method for reducing loads and generation was found to successfully remove overloads of transmission lines, ensuring stability and continuity of power. The method was found to be effective for power system protection. The method may not be optimal, but may be practical for implementation.**

Keywords—contingency, generation rescheduling, load shedding, overload, power system protection

I. INTRODUCTION

Among the last unsolved remaining problems with the power grid are blackouts, most notable of which are the 2003 blackout of Eastern US, and the 2012 blackout of India. This paper focuses on the possibility of blackout from outage of a transmission line.

When a transmission line gets disconnected, its power flow may be redistributed to the rest of the grid, which may result in overloading of remaining lines, possibly leading to cascading outages and system blackout. It is important to find effective measures to stop cascading outages arising from outage of a transmission line.

In case of an overloaded transmission line, the human operator in a control center may try to bring the current through the line within operable limits, before the transmission line gets disconnected. This may be difficult for an operator since he must consider a range of complex factors before deciding on a remedial action. In addition, there may not be enough time for human response. Clearly, automated decision-making, or helping the operator to make an informed decision are preferred.

Generator rescheduling (power reduction) and load shedding (disconnection) are two methods that can reduce overload. The overload of a single transmission line is not affected equally by all loads and generators. For removing overloads of any transmission line, it is desirable to select the most appropriate generator/s and/or load/s. Generator rescheduling or load shedding on their own may not be enough to solve overloading, and a combination of both may be more effective.

A. Literature Review

Reference [1] proposes alleviation of transmission line overload by power flow tracking, using the superposition method and proportional distribution principle. Based on the calculations, the generator output and/or load have been reduced. Reference [2] proposes alleviation of transmission line overload by decentralized communication and computation. Intelligent agents are installed at each substation for making autonomous decisions based on incomplete system information. Reference [3] proposes alleviation of transmission line overload using Particle Swarm Optimization (PSO) for generator rescheduling and/or load shedding. During a contingency, participating generators and load buses are selected using a Direct Acyclic Graph technique. Reference [3] classifies generators into a Generator Decrease group which contribute to the overload, and a Generator Increase group which do not contribute to the overload. Reference [4] proposes using local optimization for alleviation of transmission line overload by generation rescheduling and load shedding. Reference [5] proposes alleviation of transmission line overload by a knowledge-based system considering multiple transmission lines. Reference [6] proposes alleviation of transmission line overload using Newton-Raphson and fast decoupled load flow model.

B. Method of Load Shedding and Generation Reduction

This paper proposes the Reduction Method of Load and Generator Output (RMLGO) for alleviating transmission line overload. This new method determines the correlation of specific generators and loads with the overloaded transmission line. In this study, generator rescheduling mostly refers to reduction of generation output. Scheduled generation voltages mostly implies planned generator voltages.

Using the PSS®E software, it was found that transmission line outage mostly leads to low power factors of generators. The proposed method attempts to bring the power factor of the generators within acceptable limits. This method is flexible as it can incorporate other safety measures. With all these new possibilities, power system security can be further strengthened.

The proposed method was implemented using PSS®E software for power flow calculations, and was found to effectively simulate real-life power systems. It simulated

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

overloads created from different single transmission line outages on a IEEE 10-machine 39-node power system. It was successful in alleviating most overloads.

The paper is organized with Introduction in Section I. Section II has Methodology, details of transmission line overload, system constraints, the proposed method and case study. Section III provides the simulation results. Section IV has further discussion, and Section V provides the Conclusion.

II. METHODOLOGY

A. Transmission Line Overload Alleviation

Line flow can be changed by changing state variables of bus voltage magnitude and phase, which in turn can be changed by generator rescheduling and load variation [6].

Transmission lines can be overloaded even though the bus voltages are within acceptable limits. It may be effective to decrease loads at specific buses to eliminate transmission line overloads. Load shedding at receiving end buses up to two lines away from the overloaded line was found to be satisfactory [4]. Removing the load completely at the receiving end of an overloaded transmission line does not necessarily remove the overload. Overloads of specific transmission lines are affected by specific generators and specific loads. It is necessary to identify the generators and loads having the greatest influence on overloaded lines, and reduce power output of those generators or reduce those loads until there is no overloaded line.

B. PSS®E Software

Power System Simulator for Engineering (PSS®E) is a software used for high-performance transmission planning and analysis [7]. It allows analysis of networks, and is found to be reliable and accurate. PSS®E is used in industry, is popular in the developed world, and is found in the academic literature. Application Programming Interfaces (APIs) of PSS®E are based on open Python™ technology. PSS®E Xplore edition has been used, which is a free version available online.

PSS®E has been used in this study for power flow calculations and simulation of the IEEE 39 bus power system. The software uses Full Newton-Raphson power flow, to provide voltage, reactive power, and other parameters of the system. Based on the results, loads have been identified for disconnection, and generators have been identified for reducing power. Selected loads and generator outputs may be decreased simultaneously.

C. Assumptions

The following assumptions are made during power flow calculation:

- All loads have constant impedance.
- The transformer taps are fixed at the same position all the time.
- Generator terminal voltages are within limits

D. Overloaded Transmission Lines

The current flowing through the transmission line is I_{pq}. The current flowing between the buses p and q can be written as [8]:

$$I_{pq} = -Y_{pq}(V_p - V_q), \qquad p \neq q,\ p,q \in S_N \tag{1}$$

In equation (1), V_p is voltage at bus p, Y_{pq} is admittance of line between p and q, and S_N is the set of nodes or buses.

The static rating (maximum current rating) of transmission line (p - q) is $I_{pq,max}$. The transmission line overload, ΔI_{pq} is given as follows:

$$\Delta I_{pq} = I_{pq} - I_{pq,max} \tag{2}$$

The line overload alleviation problem can be considered as an optimization problem and can be formulated as below [4]:

$$Minimize\ TOL\ (X) = \Sigma_{pq \in OL} (\Delta I_{pq})^2 \tag{3}$$

In equation (3), TOL (X) is Total Overload, OL is set of Overloaded transmission lines and X is state vector of the power system, consisting of the voltage magnitudes and phase angles

To reduce the current from I_{pq} to $I_{pq,max}$ new values are needed for the variables: (i) phase angles at buses p and q (δ_p, δ_q) and (ii) Voltage magnitudes at buses p and q (V_p , V_q). In order to get the phase angles and voltage magnitudes, the power flow equations must consider different control variables. The control variables are the scheduled voltages of generators, real power output of generators, and loads at different buses.

To solve equation (3), an iterative solution is necessary since it is a large-scale nonlinear optimization problem.

E. Control Variables Selection

In the proposed method, the control variables are the loads at specific buses and scheduled voltages and power outputs at specific generators. The method shows how the values of control variables are determined whenever there is any overload. The control variables have their effect in reducing overloads.

F. Equality Constraints

The real power of the power system must fulfill the following equality constraint [3]:

$$g(x) = 0 \cong \Sigma_{i=1}^{NG} (P_{gi}) - P_d - P_L = 0 \tag{4}$$

In equation (4), P_{gi} is real power generated by the generator 'i', P_d & P_L are total real power demand and loss of the system and NG is the number of Generators.

The reactive power of the power system must fulfill the following equality constraint:

$$h(x) = 0 \cong \Sigma_{i=1}^{NG} (Q_{gi}) - Q_d - Q_L = 0 \tag{5}$$

In equation (5), Q_{gi} = reactive power generated by the generator 'i', Q_d & Q_L is total reactive power demand and loss of the system.

G. Inequality Constraints of Control Variables

The generator's active power and reactive power output needs to be within the following range:

$$P_{gi,min} \leq P_{gi} \leq P_{gi,max} \quad (6)$$

$$Q_{gi,min} \leq Q_{gi} \leq Q_{gi,max} \quad (7)$$

In equation (6), $P_{gi,min}$ is the lower limit of the active power output of generator 'i' and $P_{gi,max}$ is the upper limit of the active power output of generator 'i'. In equation (7), $Q_{Gi,min}$ is the lower limit of the reactive power output of generator 'i' and $Q_{Gi,max}$ is the upper limit of the reactive power output of generator 'i'.

H. Inequality Safety Constraints of AC Power Grid

The node voltage magnitude in pu needs to be within the following range:

$$0.95 \leq V_p \leq 1.05 \quad (8)$$

In equation (8), V_p is voltage magnitude in pu at node p.

The transmission lines need to operate within their static thermal rating:

$$-I_{pq\,max} \leq I_{pq} \leq I_{pq\,max} \; p,q \in S_N \quad (9)$$

In equation (9), $I_{pq\ max}$ is Static thermal rating of transmission line connecting the buses p and q.

The phase angle difference in degrees between two nodes p and q of a transmission line needs to be within the following range to ensure frequency is within acceptable range [9]:

$$-45^o \leq \Theta_{pq} \leq 45^o \quad (10)$$

In equation (10), Θ_{pq} is phase angle difference between nodes p and q in degrees

I. Proposed Method

The problem defined by equation (3) and all other constraints is a large-scale nonlinear optimization problem. The proposed method coordinates reduction in output of selected generators and reduction in load at selected buses. The method also fixes low power factors of generators by changing their respective scheduled voltages and by shedding reactive loads.

Initially, the safe operating currents of the different transmission lines must be provided. This enables the system to detect overloading of transmission lines.

Reduction in loads, generation output and scheduled voltages are the control actions necessary for overload alleviation.

The method is first simulated to obtain the necessary control actions. The results are then to be applied in real life.

The proposed method for power system protection after outage of transmission line is divided into 5 parts. Part 5 deals with alleviation of overload. The steps in the proposed method are given below:

Part1: Alleviation of Generator Overload

Step 1: If real power demand of generator at slack bus does not exceed maximum capacity go to Step3.

Step 2: Reduce real load at all buses until the real power demand comes within capacity.

Step 3: If reactive power demand of generators at plant buses does not exceed maximum capacity go to Step5.

Step 4: Reduce reactive load at all load buses until the reactive power demand comes within capacity. Go to Step 1.

Part 2: Alleviation of low power factor of Generators

Step 5: If the power factors of all the generators are in acceptable range, go to Step 8.

Step 6: Reduce scheduled voltages of generators having low power factors. The minimum allowable per unit voltage is 0.955. It was found that below 0.955 pu, under-voltages result at different buses. If the power factors of all generators do not come into allowable range with scheduled voltages at minimum limit, go to Step 7, or else go to Step 1.

Step 7: Reduce reactive loads in the whole system until there is no low power factor of generators. Go to Step 1

Part3: Alleviating negative reactive power of generators

Step 8: If generators are not producing negative reactive power go to Step 10.

Step 9: Increase scheduled voltages of generators at plant buses with negative reactive power generation until there is no negative reactive power. Go to Step 1.

Part 4: Voltage control at buses

Step 10. If there is no undervoltage at buses go to Step 12.

Step 11. Increase scheduled voltages of closest generator buses to the undervoltage buses until there is not undervoltage. Go to Step 1.

*Step 12.*If there is no overvoltage at buses go to Step 14

Step 13 Decrease scheduled voltages of closest generator buses to the overvoltage buses until there is no overvoltage. Go to Step 1.

Part 5: Alleviation of Transmission Line Overload

Step 14: If there is no overload terminate program.

Step 15: Reduce loads at specific buses to relieve transmission line overloads. If there are still more transmission line overloads, go to Step 16, or else go to Step 1.

Step 16: Reduce power output of specific generators to relieve transmission line overloads. Go to Step 1.

In between all steps, it is checked whether all the constraints discussed in previous sections are fulfilled.

The method for Step 14 is given below:

1) Record initial state of all loads
2) Detect overloaded branches (Branch Group A)

3) For each overloaded branches:

a) Find receiving end bus (Bus A)

b) Find receiving end bus (Bus B) connected with Bus A

c) Find receiving end bus (Bus C) connected with Bus B

d) Find out load buses from Bus A, Bus B and Bus C

4) Decrease every load by 1% in load buses found in 2.

5) Identify the branches (Identified Branch Group) from Branch Group A for which overload has decreased by more than or equal to 1%

6) Reduce the loads of corresponding load buses of Identified Branch Group from initial state until the identified branches are no longer overloaded.

The method for Step 15 is given below:

1) Record initial state of all generator outputs

2) Detect overloaded branches (Branch Category A)

3) For each generator:

a) Reduce generator output by 30% while keeping other generator outputs at 100%.

b) Identify branches in branch category A for which overload has reduced by more than 40%. When such branch exists keep record of the generator and the identified branches for each generator.

4) Reduce output of all recorded generators from initial state until their corresponding recorded branches has been relieved of overload.

Both the methods of Step 4 and Step 5 make sure of the following when finalizing control actions:

- The severity of overloaded lines has not increased.
- No previously uncritical loaded lines have become overloaded.

The software PSS®E with Full Newton-Raphson Power Flow has been used at all times.

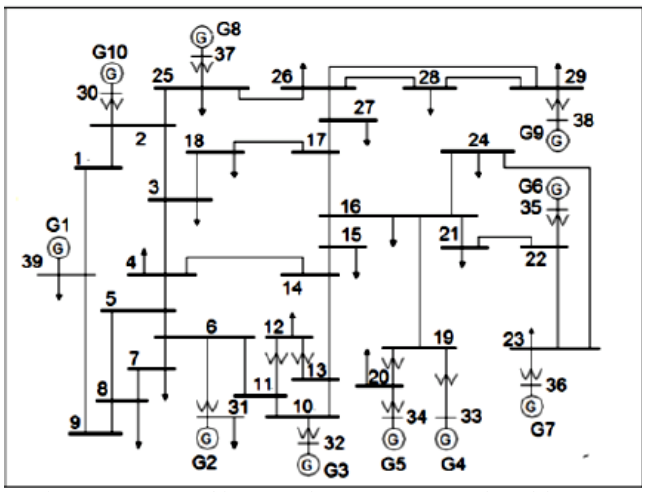


Fig. 1: IEEE 10-machine 39-node system Source: Adapted from [10]

J. *Case Study*

The proposed method was applied on the IEEE 39-bus system (fig. 1), consisting of 39 nodes, 10 generators, 34 lines and 12 transformers. The details of the test system have been found in the references [10] and [11]. Data which are not available have been assumed, and some data has been modified.

Simulation was done with the software PSS®E, Excel and Python. It is assumed that all generators and loads can take part in overload alleviation. Single transmission line outage was simulated and analyzed before, during, and after application of the proposed method.

The following three contingencies have been simulated and the output of the proposed method has been analyzed:

i. Outage of branch 25-26
ii. Outage of branch 05-06
iii. Outage of branch 13-14

III. Results of Simulation

A. *Outage of branch 25-26*

Outage of branch 25-26 leads to reduction in power factor of the generator at bus 30, and overload of branch 17-27, with current of 120% of its static rating.

The proposed method successfully brings the power factor of the generator into normal range and removes the overload. The scheduled voltages of generators 30, 31, 32 and 35 are reduced. 8% of load at bus 27 has been decreased to remove overload. The swing bus is not overloaded. The power system is in stable state after the method has been implemented.

B. *Outage of branch 05-06*

Outage of branch 05-06 leads to reduction in power factors of the generators at buses 30, 31 and 32, and overloads of branches 04-14, 06-07, 07-08, 10-13 and 13-14. The branches 04-14, 06-07, 07-08, 10-13 and 13-14 have currents of 104%, 117%, 180%, 103% and 105% of their maximum ratings respectively.

The proposed method successfully brings the power factors of the generators into normal range and also removes the overloads. The scheduled voltages of generators 30, 31 and 32 are reduced. Also total system reactive load has been decreased by 2% to bring the power factors within acceptable range. A further 4%, 7%, 19%, 3% and 19% of loads at buses 4, 7, 8, 15 and 39 has been decreased respectively to remove overloads. The swing bus was not overloaded. The power system is in a stable state after the method was implemented.

C. *Outage of branch 13-14*

Outage of branch 13-14 leads to reduction in power factors of the generators at buses 30, 31 and 32 and overloads of branches 04-05, 05-06, 06-11 and 10-11. The branches 04-05, 05-06, 06-11 and 10-11 have currents of 162%, 104%, 123% and 120% of their maximum ratings respectively.

The proposed method successfully brings the power factors of the generators into normal range and also removes

the overloads. The scheduled voltages of generators 30, 31 and 32 are reduced. Also total system reactive load has been decreased by 10%, to bring the power factors within acceptable range. A further 36% of load at bus 4 has been decreased to remove overload of the branches 04-05 and 05-06. The real power output of the generator at bus 32 was reduced by 18% to remove the remaining overloads. The swing bus was not overloaded. The power system was in stable state after the method was implemented.

IV. Discussion

A. Verification of the Proposed Method

The proposed method has been applied to different scenarios in previous sections. The results demonstrate effectiveness at alleviating transmission line overload for various scenarios. The method prevents excessive power flow throughout the grid in all the cases. It coordinates load shedding from specific nodes and generation reduction at certain nodes. This stops further spread of abnormal operation and cascading tripping.

B. Comparison with other methods

In reference [1], the power flow calculation involves only generator power and load power, which makes it insufficient when other electrical components or control variables are considered. Our proposed method can be applied to any type of power flow and is easier to implement in practice.

Reference [1] only considers active power, but reactive power should be considered in practical scenarios.

In reference [2], a decentralized approach has been taken. However, our study shows that in order to overcome overload, a centralized approach is necessary. The sensitivity of any generator and load with overloaded transmission lines vary according to the situation. Reference [2] does not include this, and considers reduction of any generator output power or any load shedding. In practice, these corrective steps may worsen the situation. Our proposed method can determine the specific generators and loads affecting a particular transmission line overload.

In reference [3], pre- and post-contingency generator areas are considered. However, it may be unsuitable when an area with multiple generators is studied, as power flow may vary due to the outage of transmission lines. Our method is not dependent on generator areas.

In reference [4], only real power is considered. But it is necessary to consider reactive power when there is transmission line outage. Our method considers reactive power and brings the power factor of generators within operable limits.

In reference [5], generators are rescheduled based on sensitivity, but it does not explain how sensitivity is measured. The sensitivity of different generators has been evaluated which has been discussed in our method.

In reference [6], the method relieves swing bus overload by shedding load at the swing bus or connected receiving end buses. However, this method does not always relieve the overload of the swing bus. In our method, it was found that curtailing certain percentage of load from the buses of the whole system is satisfactory.

C. Extending Method to a Centralized system

Full utilization of the method may be helped with a centralized system. The model of the practical power system may be created and stored in the centralized system. The power system model in the software may be updated whenever there is any change in the real-life power system.

The local information at the nodes regarding bus, load, power plant, generator, transmission line, transformer will be sent to the control center, based on which, power flow is calculated at time intervals.

The nodes should be installed with equipment to detect overloads in the transmission line and inform the control center in case of an overload. An alarm message should be added to the information when overload is detected. After the alarm message, the proposed method should be implemented at the control center.

V. Conclusion

In spite of great advances in technology, the issue of cascading failures, and blackouts are yet to be solved for the power grid [12,13].

The proposed response to transmission line overload prevents further transmission line outage. Using the software PSS®E, the method was used to coordinate a reduction in different loads and different generator outputs. This reduced overload bringing the operation of transmission lines within the permissible range. The simulation was conducted effectively on the IEEE 39 bus system.

Our method is able to identify loads and generators which have the most effect on transmission line overloads. Our method avoids needless reduction of loads and generation at other buses.

In future, Python can be used to create an interface between PSS®E software and SCADA.

The proposed method can be used for renewable energy sources, and may be effective even if islanding has taken place.

The proposed method is not claimed to be an optimal solution; rather it provides a feasible solution in the time of emergency.

Further studies can be conducted as follows:

(1) Our method has the potential to reduce overload arising from different types of contingencies. It may be evaluated for different types of N-1 contingencies and for beyond N-1 contingencies.

(2) Our method may be evaluated for different types of loading scenarios and different weather conditions.

(3) Our method may be tested with other corrective controls such as phase shift adjustment, dc tap adjustment, switched shunt adjustment, and optimal power flow.

References

[1] D. Jiandong, C. Wenji and X. Bing, "Line Overload Shedding Strategy Based on Improved Power Flow Tracking Algorithm," *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, Xi'an, China, 2019, pp. 2315-2320.

[2] L. Robitzky, S. C. Müller, S. Dalhues, U. Häger and C. Rehtanz, "Agent-based redispatch for real-time overload relief in electrical transmission systems," *2015 IEEE Power & Energy Society General Meeting*, Denver, CO, 2015, pp. 1-5.

[3] M. K. Maharana and K. S. Swarup, "Particle Swarm Optimization based corrective strategy to alleviate overloads in power system," *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)*, Coimbatore, 2009, pp. 37-42.

[4] A. Shandilya, H. Gupta and J. Sharma, "Method for generation rescheduling and load shedding to alleviate line overloads using local optimisation," in *IEE Proceedings C - Generation, Transmission and Distribution*, vol. 140, no. 5, pp. 337-342, Sept. 1993.

[5] J. S. Chen, E. D. Tweed and E. F. Richards, "A knowledge-based system to develop a plan of action for eliminating transmission line overloads," *The Proceedings of the Twenty-First Annual North American Power Symposium*, Rolla, MO, USA, 1989, pp. 21-32.

[6] T. K. P. Medicherla, R. Billinton and M. S. Sachdev, "Generation Rescheduling and Load Shedding to Alleviate Line Overloads-Analysis," in *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, no. 6, pp. 1876-1884, Nov. 1979.

[7] Siemens, 'PSS®E, High performance transmission planning and analysis software' , 2020. www.pss-store.siemens.com [Accessed: 13- Apr- 2020].

[8] A. Ghosh, Class Lecture, Topic: "Chapter 4: Load Flow Studies." NPTEL, Chennai, India, April, 2020

[9] J. Hu, J. Wang, X. Xiong and J. Chen, "A post-contingency power flow control strategy for AC/DC hybrid power grid considering the dynamic electrothermal effects of transmission lines," in *IEEE Access,* Vol. 7, pp. 65288-65302, Jun 2019.

[10] A. Agatep, "Voltage stability analysis using simulated synchrophasor measurements," M.S. Thesis, California Polytechnic State University, California, USA, 2013.

[11] Y. Zhang, "Design of wide-area damping control systems for power system low-frequency inter-area oscillations," Ph. D thesis, School of Electrical Eng. and Computer Sc., Washington State University, Washington, USA, 2007.

[12] K. M. J. Rahman, M. M. Munnee and S. Khan, "Largest Blackouts Around the World: Trends and Data Analyses", *IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE),* pp. 155-159 (2016)

[13] S. Khan, Power System Operation and Protection, S. Khan, 2nd Edition, Dhaka, Bangladesh (2020)

Parameter estimation and its application in tuning PI control scheme

Sanjay Bhadra
Department of Electrical Engineering
University of Engineering and Management
Kolkata, West Bengal, India
sanjay.bhadra@uem.edu.in

Atanu Panda
Department of Electronics and Communication Engineering
Institute of Engineering and Management
Kolkata, West Bengal, India
atanu.panda@iemcal.com

Sawan Sen
Department of Electrical Engineering
Kalyani Government Engineering College
Nadia, West Bengal, India
sawansen@gmail.com

Parijat Bhowmick
Department of Electrical and Electronic Engineering
University of Manchester
Manchester, UK.
parijat.bhowmick@manchester.ac.uk

***Abstract*— In this paper, Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF) based robust adaptive PI Control Scheme have been proposed and implemented on the simulated model of the non-linear benchmark processes. The servo and regulatory performance using proposed tuning methodologies were found satisfactory. The performances of the proposed control schemes have been compared with conventional adaptive PI (CA-PI) control scheme. From the extensive simulation studies, it was found that the proposed schemes implemented on non-linear processes are having better performance over CA-PI control scheme. It was also found that proposed control schemes are able to eliminate measurement noise and having good robustness features.**

Keywords—Model-based control, parameter estimation, UKF, EKF, spherical tank process, tracking.

I. INTRODUCTION

This template P/PI/PID and model based control schemes are the widely used control schemes over past few decades. Different kinds of adaptive control schemes such as gain scheduled adaptive controller, self-tuning controller and model reference adaptive controller are widely-used control-techniques for various types of systems whose parameters vary with respect to time [2]. The design plethora of deriving tuning methodologies for conventional adaptive PI control schemes (CA-PI) on the basis of different performances have been discussed in the literature [2-3].

As a well-recognized example of feedback control loop, CA-PI scheme was improvised for non linear processes. However, CA-PI control scheme specifies output feedback to formulate driving action for generating proper manipulated variable to control the system output [2]. The noteworthy phenomenon of CA-PI control technique deals with tradeoff between plant-set point mismatch [13]. Several articles have already been recognized about different ways of approaches for CA-PI control schemes like direct synthesis mehod [10,19], Skogestad's method [1], Ziegler-Nichol's method [19], Cohen-Coon's method [13], Hagglund-Astrom's method [3], Tyreus-Luyben's method [4], Relay method [19], Gain-Scheduling method [18], Setpoint overshoot method [19], Good Gain method [19], IMC-PID method [13] and so on. In general, PID control scheme can be classified in four categories namely PID0, PID1, PID2 and PID3 [13]. In case of PI control scheme both PID1 and PID2 control strategies are equivalent and our proposed work comes under this category. Furthermore, the needing of adaptation mechanism in PI control techniques has been explained by Michael and Costas (2003), which shows significant improvement in feedback control performances.

Thus, the objective of this simulation study employs implementation of ydifferent tuning methodologies in PI control scheme for chemical processes. A typical structural characteristics based models have been studied and considered in this proposed work. The motivation of the proposed work lead to developing an accurate inexpensive, less computationally hazardous controller which would be able to track desired value and would be able to eliminate load disturbances. Performance of the proposed work shows significant improvement in set point tracking or uncertainties associated with process/load disturbance.

Maintaining In this work, a comparison study has been made between proposed works with CA-PI control technique. From the extensive simulation study an illustration of performances like ISE (integral square error) and TV (total variation) for the mentioned control schemes have been calculated. The robustness of all the mentioned controllers also have been discussed in this study. Best of our knowledge, the proposed work has not been addressed in the literature.

II. CONTROL SCHEMES

A. Conventional adaptive PI control scheme (CA-PI):

The adaptive conventional PI control law has been described by the general equation as [3,14,21]:

$$u(k) = u(k-1) + (k_{c1} * \partial e_1(k)) + (T_s * (k_{c1} / T_i) * e(k)) \quad (1)$$

T_s is the sampling time. e and ∂e are the process error and change in process error and can be expressed as:

$$e(k) = y_{sp}(k) - y(k) \quad (2)$$

$$\partial e(k) = e(k) - e(k-1) \quad (3)$$

Where $(k_{c1}; T_i)$ is the conventional adaptive PI controller tuning parameters. We have named conventional adaptive PI controller as CA-PI shortly.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

B. Proposed adaptive PI control scheme using UKF tuning methodology:

In this case, adaptive PI controller tuning parameters have been estimated by UKF algorithm. The system equations for the controller parameters of the adaptive PI controller are as follows [9,14-15]:

$$\theta(k) = \theta(k-1) + w(k-1) \tag{4}$$

$$\theta(k) = [k_{c1}; T_i] \tag{5}$$

$$pv(k) = f[pv(k-1), u(k-1), \theta(k-1)] + v(k) \tag{6}$$

pv is the measured value (of process variable) obtained from the plant. The predicted controller parameters estimation are obtained as

$$\hat{\theta}(k \mid k-1) = \hat{\theta}(k-1 \mid k-1) \tag{7}$$

The covariance matrix of estimation errors in the predicted controller parameters estimation are obtained as follows:

$$P_\theta(k \mid k-1) = P_\theta(k \mid k-1) + Q \tag{8}$$

A set of $(2L+1)$ sigma points with the associated weights $w(i)$ are chosen symmetrically about $\hat{\theta}(k \mid k-1)$ as follows:

$$\theta_s = [\hat{\theta}(k|k-1), \hat{\theta}(k|k-1) + \sqrt{(L+\kappa)P_\theta(k|k-1)}, \hat{\theta}(k|k-1) - \sqrt{(L+\kappa)P_\theta(k|k-1)}] \tag{9}$$

The measurement prediction $\hat{pv}(k \mid k-1)$, computation of innovation $(e_{k|k-1})$ covariance matrix of innovation $(P_{ee}(k))$ the cross covariance matrix between the predicted model parameter estimation errors and innovation $(P_{\theta e}(k))$ are computed as follows:

$$\hat{pv}^i(k|k-1) = f(pv(k-1), u(k-1), \theta^i(k|k-1)) \tag{10}$$

$$\hat{pv}(k \mid k-1) = \sum_0^{2L} \hat{pv}(k \mid k-1) \tag{11}$$

$$e_{k|k-1} = \hat{pv}(k) - pv(k) \tag{12}$$

$$P_{ee}(k) = \sum_{i=0}^{2L} w_i (\hat{pv}^i(k|k-1) - pv_r(k|k-1))(pv_r^i(k|k-1) - pv_r(k|k-1))^T + R \tag{13}$$

$$P_{\theta e}(k) = \sum_{i=0}^{2L} w_i (\hat{\theta}_s^i(k|k-1) - \hat{\theta}(k|k-1))(\hat{pv}^i(k|k-1) - pv_r(k|k-1))^T \tag{14}$$

Where $w_0 = \kappa/(L+\kappa)$

$w_i = \kappa/(2(L+\kappa))$

The Kalman gain is computed using the equation

$$K = P_{\theta e}(P_{ee}^{-1}) \tag{15}$$

The updated model parameter estimates are obtained using the equation:

$$\hat{\theta}(k \mid k) = \hat{\theta}(k \mid k-1) + Ke_{k|k-1} \tag{16}$$

The covariance matrix of estimation errors in the updated model parameter estimation are obtained as:

$$P_\theta(k \mid k) = P_\theta(k \mid k-1) - KP_{ee}K^T \tag{17}$$

C. Proposed adaptive PI control scheme using EKF tuning methodology:

In this case, adaptive PI controller tuning parameters $(k_{c1}; T_i)$ have been estimated by EKF algorithm. The system equations for the controller parameters of the adaptive PI controller are same like equation [1-3].The predicated controller parameters'estimation areobtained as:

$$\hat{\theta}(k \mid k-1) = I * \hat{\theta}(k-1 \mid k-1) \tag{18}$$

The covariance matrix of estimation errors in the predicted controller parameters estimation are obtained as follows:

$$P_\theta(k \mid k-1) = \varphi * P_\theta(k-1 \mid k-1) * \varphi^T + Q \tag{19}$$

where φ is the phase transition matrix. Measurement covariance can be obtained as:

$$S_k = R + H_j * P_\theta(k \mid k-1) * H_j^T \tag{20}$$

H_j is the Jacobian of the observation matrix. The Kalman gain is computed using the equation:

$$K = P_\theta(k \mid k-1) * H_j^T * (S_k)^{-1} \tag{21}$$

The updated model parameter estimates are obtained using the equation:

$$\hat{\theta}(k \mid k) = \hat{\theta}(k \mid k-1) + Ke_{k|k-1} \tag{22}$$

$$P_\theta(k \mid k) = P_\theta(k \mid k-1) - K * H_j * P_\theta(k \mid k-1) \tag{23}$$

In this proposed work a performance comparison study (ISE and TV) has been made between proposed control schemes with conventional adaptive PI control technique. The ISE and TV can be derived as follows [18]:

$$ISE = \int_0^\tau (y_{sp}(k) - y(k))^2 dk \tag{24}$$

$$TV = \sum_{i=1}^{k} |u(k) - u(k-1)| \tag{25}$$

III. Process Description

The processes considered for the simulation study is spherical tank, CSTR and pH processes.

A. Spherical tank system: The material balance equation for the spherical tank system is as follows [9,20]:

$$A(h)\frac{dh}{dt} = f_{in} - f_{out} \tag{26}$$

Where A(h) is the area of the tank. f_{in}(cc/sec) and f_{out} (cc/sec) are inflow and outflow rate of the tank respectively. H(cm) is the water level in the spherical tank system. The relationship between outflow rate and the water level (h) in the tank is as follows:

$$f_{out} = c_v \sqrt{2gh} \tag{27}$$

Where c_v is the discharge co-efficient. It should be noted that the area of the spherical tank varies with water level in the tank. In case of spherical tank, the area of the tank can be calculated using following relations.

$$A(h) = \pi \left[2rh - h^2 \right] \tag{28}$$

Where r is the maximum radius of spherical tank.

Table I. Values of the process parameters

Process variable	Nominal operating value
Maximum Height (H)	50 cm
Maximum Radius (R)	25 cm
Valve Coefficient (C_v)	0.5

All the simulations were executed by considering first principle model mentioned in equation (26-28). True state can be obtained by solving differential equation using MATLAB 7.2 toolboxes. In the entire simulation studies, sampling time has been considered as 0.0833 min. A constrained on the manipulated variable $(0.01 < f_{in} < 20cc/\sec)$ has been imposed. Following operating point has been taken for entire simulation studies $(\bar{h} = 30, f_{in} = 2.7386)$. The controller performances (ISE and TV computation) of all mentioned control schemes at servo level, regulatory level and in presence of measurement noise have been reported in Table II.

B. Open loop study: In order to assess the open loop study of the spherical tank system, a sequence of step changes (combination of positive and negative steps) in the manipulated variable has been introduced. Fig. 1(a) represents the steps changes in the input (f_{in}). The variation of process output (h) is reported in Fig 1(b).

C. Servo response:

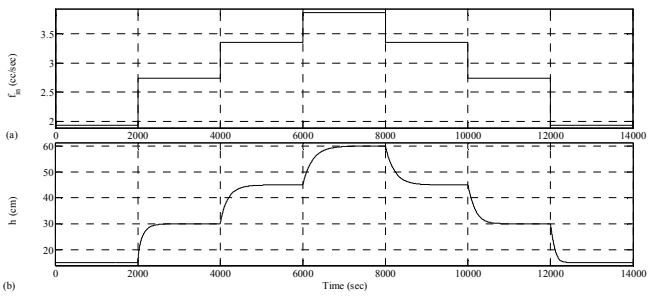


Fig. 1 Open loop response of the spherical tank processes (a) variation of input, (b) process output.

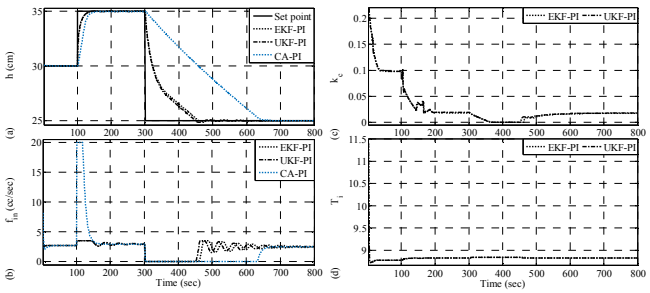


Fig. 2 Servo response of the spherical tank system with various tuning methodologies - (a) process output, (b) variation of controller output, (c) k_c, (d) τ_i.

In order to assess the tracking capability of all the above mentioned control schemes, set point variation as shown in Fig. 2(a) has been introduced. Simulation studies were made based on the set point variation of 30 to 35 and finally 35 to 25. From Fig. 2(a), it can be inferred that the all the controllers are able to maintain the set point at desired level. The variation of controller output is reported in Fig. 2(b). The evolution of the EKF-PI, and UKF-PI controller tuning parameters $(k_{c1}; T_i)$ have been shown in Fig. 2(c) and Fig. 2(d) respectively. From the simulation study, it was observed that CA-PI controller performs aggressively at the time of set point change. However, it can be stated that robustness of both EKF-PI and UKF-PI controllers were found satisfactory and better over CA-PI control scheme.

D. Regulatory response:

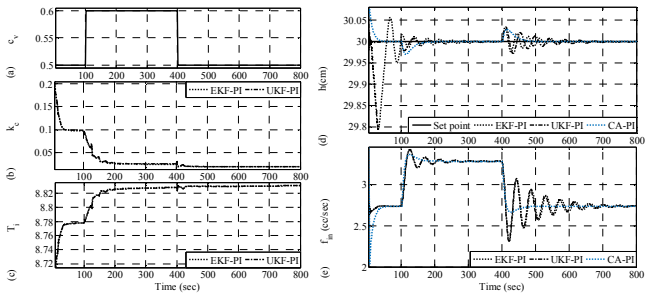


Fig. 3 Regulatory response of the spherical tank system with various tuning methodologies - (a) Evolution of downstream valve (c_v), (b) k_c, (c) τ_i, (d) process output, (e) variation of controller output.

In order to assess the disturbance rejection capability of all the mentioned control schemes, step like changes in the downstream valve position as shown in Fig. 3(a) has been introduced. In this simulation study, set point was maintained at 30 cm. It should be noted that a positive step change in the downstream valve position of magnitude 0.05 (from 0.5 to 0.55) has been introduced at 101th sampling instance and the same value has been maintained up to 400th and again negative step change in the downstream valve position of magnitude 0.05 (from 0.55 to 0.5) has been introduced at401th sampling instance and the same value has been maintained up to 800th sampling instance. From Fig. 3(d), it can be inferred that the all the controllers are able to reject disturbance and

bring back process variable to the desired level. The variations of controller outputs are reported in Fig. 3(e). The evolution of the EKF-PI, and UKF-PI controller tuning parameters (k_{c1}, T_i) have been shown in Fig. 3(b) and Fig. 3(c) respectively. From the simulation study, it can be concluded that eliminating disturbances with CA-PI control scheme is taking longer time to settle.

E. Measurement noise elimination:

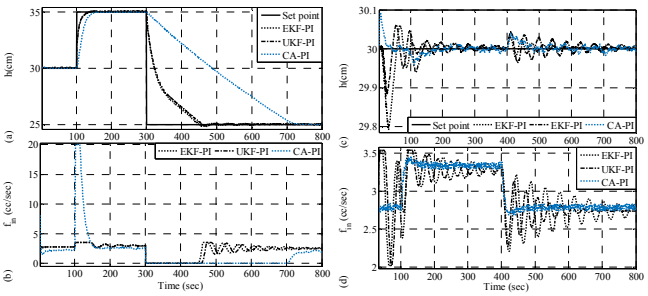


Fig. 4 Servo and regulatory response of the spherical tank process with various tuning methodologies in presence of measurement noise (NSR = 0.02) - (a) process output - servoresponse, (b) variation of controller output- servo response,(c) process output - regulatory response,
(d) variation of controller output- regulatory response.

As measurement noise is a common phenomenon in real time plant, hence efficacy of the controller can be identified by checking the ability to track at desired level in presence of measurement noise [4]. In order to introduce measurement noise into the process, white random noise has been added into the plant output. The efficiency of both EKF-PI and UKF-PI control schemes in presence of measurement noise have been measured and compared with CA-PI control scheme.

It should be noted that, for introduction of measurement noise, NSR (noise to signal ratio) has been taken as 0.02. From Fig. 4(a), it can be inferred that all the control schemes are able to eliminate measurement noise and bring back process variable to the desired value, whereas regulatory performances in presence of measurement noise was evaluated successfully (see in Fig. 4(c)). The evolution of controller outputs at servo and regulatory level are reported in Fig. 4(b) and Fig. 4(d) respectively.

F. Assesing Robustness of the Controller:

In order to assess the robustness feature for both EKF-PI and UKF-PI control schemes, changes of parameters (downstream valve co-efficient) at different point of time between plant and model have been introduced (Shown in Fig. 3(a)). From Fig. 3(d), it can be inferred that all the control schemes are able to maintain water level at desired set-point in presence of plant and model mismatch. The variations of controller outputs have been shown in Fig. 3(e). The ISE and TV values of all the control schemes in the absence and presence of process and model mismatch (10 % deviation of discharge coefficient) are reported in Table II. From Table II, it can be concluded that there are deterioration in performances for all the above control schemes in the presence of plant and model mismatch.

G. Performance measurement of different control schemes:

In order to assess performances of different control schemes for servo-regulatory response and in presence of measurement noise, ISE and TV computation have been reported. Table II represents ISE and TV computation of spherical tank process respectively. Servo performance of spherical tank was made based on set point variation as shown in Fig. 2(a), whereas regulatory performance was made based on Fig. 3(a) and Fig. 3(d). From the ISE chart for regulatory performance, it can be inferred that CA-PI scheme performs better over both EKF-PI and UKF-PI schemes, whereas EKF-PI and UKF-PI schemes performs better over CA-PI scheme in servo level. It was also noticed that there is deterioration in performances for all the control schemes in presence of measurement noise. From the performance chart, it can be concluded that, there is slight difference in performances between EKF-PI and UKF-PI, where UKF-PI performs slight better than EKF-PI scheme.

Table II. ISE chart for servo, servo-regulatory responses and in presence of measurement noise (NSR = 0.02) for various control schemes implemented on spherical tank process.

Control schemes	Servo level	Regulatory level	Servo level (with measurement noise)	Regulatory level (with measurement noise)
CA-PI	1.0793*e+04	8.7073*e-02	1.2681*e+04	1.4423*e-01
EKF-PI	1.7682*e+03	8.3566*e-01	1.9546*e+03	1.0428*e-00
UKF-PI	1.7680*e+03	8.3563*e-01	1.9544*e+03	1.0427*e-00

IV. CONCLUSION

In this paper, a successful design and implementation of EKF and UKF based adaptive PI control schemes have been discussed. From the simulation study, it can be concluded that servo and regulatory performances of EKF and UKF based PI controller (both presence and absence of measurement noise) implemented on spherical tank was found satisfactory. From the comparison study, it can be inferred that proposed control schemes take very less time to settle over CA-PI controller at the time of set point variation. It can also be concluded that proposed tuning based adaptive PI control schemes take less time to eliminate disturbances over CA-PI control scheme. It was also observed that, CA-PI control scheme is having poor robustness compared to proposed tuning based adaptive PI control schemes. From the controller performances chart (ISE and TV computation), it can be inferred that proposed controllers are having better performance over CA-PI control schemes. It was also noticed that proposed control schemes are good to eliminate measurement noise compared to CA-PI control scheme. From overall performance of all the mentioned control schemes, it can be concluded that proposed control schemes performs better over CA-PI control law.

REFERENCES

[1] Rivera D.E., Morari M. and Skogestad S., Internal model control: PID controller design, Ind.Eng.Chem.Res.25(1):252-265(1986).

[2] Bequette BW (1991) Nonlinear control of chemical processes: processes: a review. Ind Eng Chem Res.30(7):1391-1413.

[3] Astrom K.J., T. Hagglund, Hung C.C. and Ho W.K. (1992), Automatic tuning and adaptation in. PID controllers- Asurvey. IFAC Proceedings. 25(14): 371-376 .

[4] Azlan HM, Paisan Kittisupakorn and Wachira Daosud (2001) Implementation of neural-network based inverse-model control strategieson an exothermic reactor.Science Asia 27:41-50.,

[5] Tan KK, Huang S and Ferdous R, Robust self - tuning PID controller for nonlinear systems, Journal of Process Control 12(7):753-761.

[6] S. Haykin, Kalman Filtering and Neural Networks, John Wiley & Sons Inc(2001).

[7] Galan O, Josh A. Romagnoli and Ahmet Palazoglu, Real-time implementation of multi-linear model based control strategies-An application to a bench scale ph neutralization reactor. J. of Process Control,14(5):571-579(2004).

[8] Rames C Panda, Cheng-Ching Yu and Hsiao-Ping Huang (2004), PID tuning rules for SOPDT systems:review and some new results, ISA Transactions 43(2):283-295.

[9] Michael W. Foley, Navin R. Ramharack and Brian R. Copeland, Comparison of PI controller tuning methods, Ind. Eng. Chem. Res, 44(17):6741-6750(2005).

[10] K. K. Ahn and D. Q. Truong, Online tuning fuzzy PID controller using robust extended Kalman Filter, J. of Process Control 19(6):1011-1023(2009).

[11] A Panda, RC Panda, Adaptive nonlinear model-based control scheme implemented on the nonlinear processes, Nonlinear Dynamics 91(4):2735-2753.

[12] Finn Haugen (2010), Comparing PI Tuning Methods in a Real Benchmark Temperature Control System, Modelling Identification and control 31(3):79-91(2010).

[13] S Chandrasekharan, RC Panda, BN Swaminathan, A Panda, Operational control of an integrated drum boiler of a coal fired thermal power plant, Energy 159:977-987.

[14] Mohammad Rohmanuddin and Augie Widyotriatmo, A novel method of self-tuning PID control system based on time-averaged Kalman filter gain, 3rd International Conference on Instrumentation Control and Automation(ICA), Indonesia(2013).

[15] Meysam Shadkam, Hamed Mojallali (2013), Speed Control of DC Motor Using extended Kalman Filter based Fuzzy PID, International Journal of Information and Electrincs Engineering, 3(1).

[16] Kapil Arasu S, Atanu Panda and Prakash J, Experimental validation of nonlinear model based control scheme on the variable area tank process, 9th IFAC Symposium on advanxces in control and optimization of dynamical systems(IFAC-ACDOS 2016), NIT Trichy pp 30-34 IEEE Publishing.

[17] Omer Saleem, Urwa Omer, EKF based self regulation of an adaptive nonlinear PI speed controller for a DC, Turk J Elec Eng & Comp Sci.25:4131-4141(2017).

[18] Atanu Panda and Prakash J, State estimation and nonlinear model based control of a continuous stirred tank reactor using Unscented Kalman Filter, The Canadian Journal of Chemical Engineering.95(7):1323-1331(2017).

[19] S Bhadra, A Panda, P Bhowmick, S Goswami, RC Panda, Design and application of nonlinear model based tracking control schemes employing DEKF estimation, Optimal Cont. Appl. Methods 40(5):938-960.

-

Autonomous Mobility Vehicle

Alston Brazil Coutinho
Koblenz, Germany
alstoncoutinho45@gmail.com

Rehan Ali Mirza
Mechanical, Automotive & Materials Engineering (MAME)
University of Windsor
Windsor-Essex County, Ontario, Canada
rehanali_2640@outlook.com

***Abstract* — This paper describes about the development and conversion of a mobility vehicle into an autonomous mobility vehicle using image processing, ultrasonic and infrared sensors, relays etc. The paper focuses on the rebuilding of the vehicle, interfacing the various sensors and the image processing module to work in sync with the vehicle to cause the vehicle to navigate autonomously on lanes dedicated for pedestrians without collision. The rebuilding phase involves a lot of the mechanical work done on the vehicle which includes turning the vehicle from scrap to working condition and converting the manual steering into an automatic steering by installing a motor for the same. The image processing bit involves lane detection algorithms to actuate the vehicle and pedestrian detection algorithms using YOLO object detection which involves machine learning. The paper also describes the use of a line following system in custom designed navigation lines specifically only for indoor environments where there are no lanes. (*Abstract*)**

Keywords— Autonomous; vehicle; mobility; image processing; (*key words****)***

I. INTRODUCTION

Daily Commuting can be quite hectic for old or handicapped people without another person accompanying them for their daily activities. This is a bigger challenge when an old or a handicapped person has to go for purchasing daily or weekly groceries or using public transport on a regular basis. Doing such activities will be challenging for old/handicapped people without a regular pedestrian accompanying them or at least getting help from nearby pedestrians [1].

An autonomous mobility vehicle should be capable of sensing its surroundings and navigate without any human input or intervention. It uses a range of techniques to find its surroundings, like radiolocation, odometry and computer vision. Advanced management systems interpret sensory information to spot acceptable navigational methods for avoiding obstacles. It also has management systems which are capable of analysing sensory information to differentiate between different cars and pedestrians [2].

A fully computerized automobile capable of doing nearly everything a human driver can do is the main attraction of autonomous vehicles. The various activities that can be performed by them are for example, picking up kids from school, picking up relatives from the airport, going on long drives with less fatigue on the driver etc. Our main focus is for the elderly and physically challenged to be able to commute independently. As a matter of fact, vehicles around the planet are currently equipped with intelligent devices that allow the vehicles to respond to various factors such as abrupt acceleration or braking [3].

The system utilizes image processing by using a webcam capable of at least 720p video recording and four ultrasonic sensors, in which three ultrasonic sensors would be at the front i.e. one at the right and one at the left as well as one at the center and the fourth one at the rear. The input would be taken by the sensors which would increase the accuracy of the vehicle to detect obstacles [4]. The motors are actuated by the programming device based on feedback from the camera and the respective sensors. The speed is adjusted automatically based on the feedback from the sensors, respectively.

When there is an obstacle within a minimum distance set by the user, the vehicle stops [4]. The feedback from the camera is taken and sent to the processing unit of the vehicle where lane detection and object detection is done. Lane detection module helps the vehicle to stay within a lane [5]. This is particularly useful in huge office campuses and college campuses where there are roads with lanes.

To process the video feed for image processing, we would require a processor that is capable enough to process images at a desirable framerate. The ideal processing module for this would be one of the boards from NVidia Jetson Tegra series built specifically for self-driving cars [6]. A screen would need to be attached to this which would help the user access the GUI and facilitate the user to give instructions to the vehicle. The screen must be wireless, a tablet would be ideal due to its low cost and its capacitive touch screen.

The main goal of this project is to present an autonomous mobility vehicle that has the capability to detect and follow lanes such as pedestrian lanes while simultaneously avoiding any obstacles in its path using sensors as well as detect and recognize various objects in its environment.

Along with outdoor capabilities the vehicle can also be utilized in indoor spaces such as malls or plazas with a lane line designed specifically for this vehicle. Hence, a line following system would be an added benefit.

To navigate within an indoor environment, it would need to follow lines instead of lane lines. For line following, we could

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

potentially add two I/R sensors in the front to make the vehicle continue in a line.

This paper comprises of the following sections. Section II deals with the various components involved in this project. Section III explains the interfacing the sensors and actuators with the microcontrollers. Section IV is the implementation of the lane detection algorithm. Section V puts more emphasis on the real time object detection and finally Section VI details the results and conclusion of this paper.

II. Components Involved

The control unit of the autonomous mobility vehicle includes micro-controllers and processing modules. They are as follows:

1. Arduino Mega 2560

Arduinos are micro-controllers which can control and interface most of the motors, relays and sensors such as ultrasonic sensors, infrared sensors etc. Arduino Mega has 54 I/O Pins, out of which 16 pins are PWM I/O pins. It has a USB Port which was used to communicate with the processing module. Using the module Pyfirmata, the processing module can communicate using a set of commands in python with the Arduino. Out of the 54 I/O ports about 20 ports were used to interface the other components such as relays, servo motors and sensors etc. The Processing module is also responsible for powering the Arduino, since it is connected to the processing module using the USB port.

2. Servo Motor (SG90)

The accelerator of the vehicle is controlled by a 2k potentiometer. To actuate the potentiometer, we needed an appropriate motor to actuate it. The most appropriate motor to serve this purpose is a 5v Tower Pro (SG90) servo motor.

3. Relay

Except the acceleration, the rest of the vehicle is controlled by relays. Two dual channel relays were used to control the steering and a single channel relay was used to control the forward and reverse gear of the vehicle.

4. Tie Rod

To form the steering system a rod was used. A single tie rod was to connect the window motor to the steering mechanism to control the steering. Tie rod ends are basic components that connect the steering rack to the steering knuckle on the front wheels. Once the wheel is turned, the movement is transmitted through various steering parts till the rod ends are pushed or pulled to make the wheel change direction.

5. DC Motor

The window motor (which is a DC motor) is used for controlling the steering mechanism of the vehicle. As our vehicle is an electrical vehicle, DC motor is used here. This motor is controlled by a series of relay circuits which is controlled by the Arduino.

6. Ultrasonic Sensor

For this project four ultrasonic sensors were used, three at the front and one at the back. When an obstacle is 2 meters within range of the sensors placed on the vehicle, the processor would give a signal to the propelling motors to come to halt. The sensors work in tandem with the rest of the modules of the vehicle to avoid any sort of collision. The Ultrasonic sensor which we used is HC- SR04. This sensor has a range of 2cm to 400 cm. It has four pins namely Vcc, Ground, Trigger and Echo.

III. Interfacing the Vehicle

1. Controlling the Vehicle.

For the vehicle to be autonomous one of the key features is to have a complete control over the steering mechanism. In order to gain control over the steering mechanism, we had to attach a motor to it. The right choice of motor was a DC motor, since it must handle the weight of the vehicle as well as the weight of the passenger sitting on the vehicle.

The steering motor is controlled by the signals received from the Arduino. The Arduino sends signals to the 5 volt dual channel relay, which sends signals to the 12v dual channel relays to actuate it in the desired direction.

2. Use of Relays

2.1 Steering

The 12 volt dual channel relay is connected to the motor and the power source. Two digital pins from the Arduino is connected to a dual channel 5v relay. this 5v relay is connected to a 12v dual channel relay circuit on a printed circuit board. The 12v dual channel relay circuit is connected to the motor and a 12v power source. The 12v power source is connected to the 5v relay. Upon giving a trigger signal from the Arduino Uno to one of the 5v volt relays, it'd close the circuit which would give a trigger signal to the 12v relay to actuate the motor in a particular direction. The two 12v batteries are connected in series to power the ECU, we take a parallel connection from the batteries connected in series which gives a total of 24 volts. In order to draw the required power from the batteries we use a power converter to reduce the power to 12v and 30 amps as this is the requirement of the steering motor. This is a 24 volt to 12 volt, DC to DC 30 amps power converter.

2.2 Forward Reverse Switch

The forward reverse switch was a two way switch which had to be manually actuated by the user. In order to automate this process we use a single channel 5v relay with the Power connected to the common. The Forward switch would be connected to normally closed and the Reverse Switch would be connected to normally open. So when this relay is triggered the vehicle switches to reverse and when it isn't triggered the vehicle is in forward mode by default. The trigger to this relay is given by the Arduino.

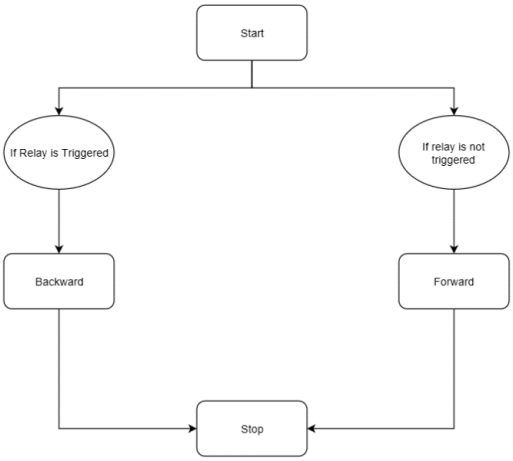


Fig. 1. Forward Reverse Flowchart

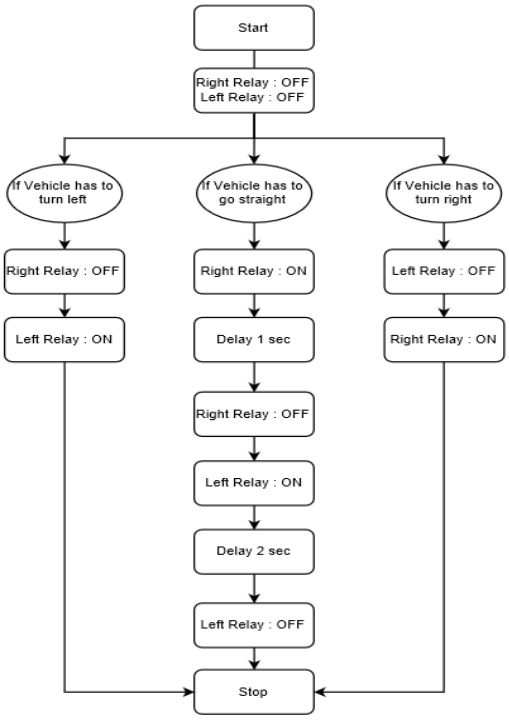


Fig. 2. Steering Flowchart

2.3. Controlling the Acceleration

The accelerator of the vehicle is controlled by a 2k potentiometer which works by varying the voltage by varying the resistance. The mechanism already in place was a spring pulley which actuate the potentiometer based on how much the shaft was rotated. In order to automate this process, we used a 5v Tower Pro SG90 servo motor to attach to the knob of the potentiometer. The servo receives signals from the Arduino for it to actuate to certain angles to control the potentiometer. The servo motor actuates accordingly to the specified angle and as a result the potentiometer moves along with the servo. This results in the acceleration of the propelling dc motor based on the varied voltage signals.

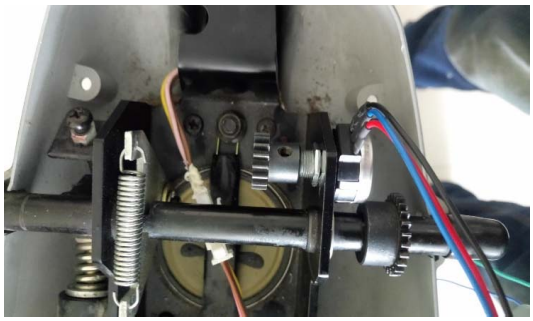


Fig. 3. Replacement of the spring mechanism with a Servo motor for the accelerator

IV. LANE DETECTION USING OPENCV

Python and OpenCV are used to find lane lines on the road and to accomplish this task we use the following techniques:

1. Colour Selection
2. Canny Edge Detection
3. Region of Interest Selection
4. Hough Transform Line Detection

1. Colour Selection:

We have our lane images in RGB (Red Green Blue) which is the normal image that we see on our screen:

Fig. 4. Video feed images of lanes

Now as we can observe in the above images, the lane lines are either yellow or white. So, the first step in detecting lane lines is to create a mask of yellow and white color that will show only these two colors and anything apart from yellow or white is rendered as black or null.

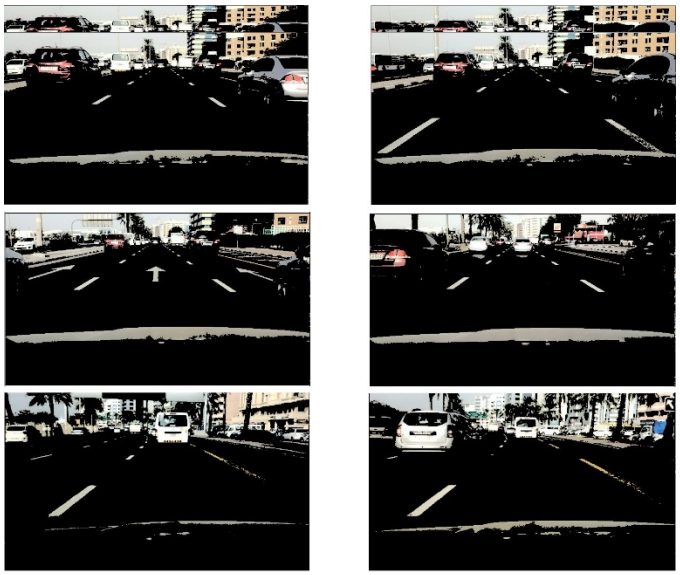


Fig. 5. Applying white and yellow color selection to HSL mask

We need to choose a different color space that does not alter the image like RGB does when there is a shadow. Therefore, we can choose either HSV or HSL color space. We went with HSL color space as it gave better results.

We can notice the yellow lanes are clearly visible in the last two images. This is because HSL color space does not change the value of the pixels based on darkness unlike RGB.

2. Canny Edge Detection:

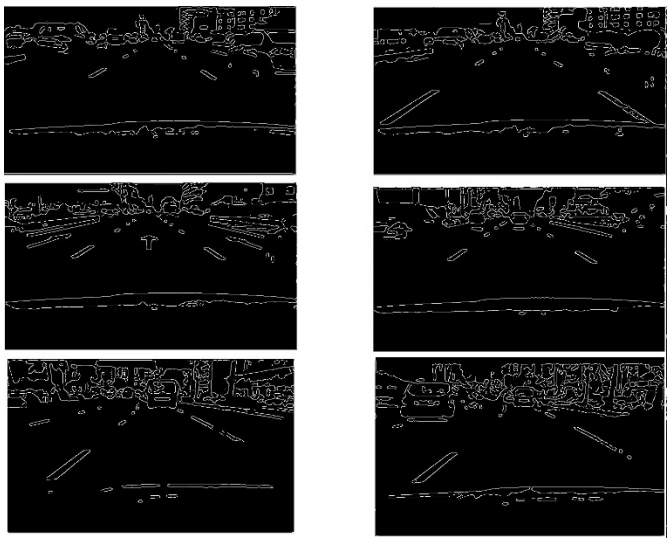


Fig. 6. Applying Canny edge detection

Since we have only the yellow and white lanes visible in the images, we want our computer to now detect them and to achieve that we first need to detect the edges of the lanes using Canny Edge Detector which is built in function in OpenCV [7]. However, we need to first prepare the image for Canny Edge detection, and this involves converting the images to Gray Scale and then applying Blurring effect and finally applying the Canny edge detector to these blurred images in order to detect the edges. The images with edge detection are as follows:

3. Region of Interest:

When finding lane lines, we do not need to check the sky and other objects in the environment. Roughly, speaking, we are interested in the area surrounded by the red lines as shown below:

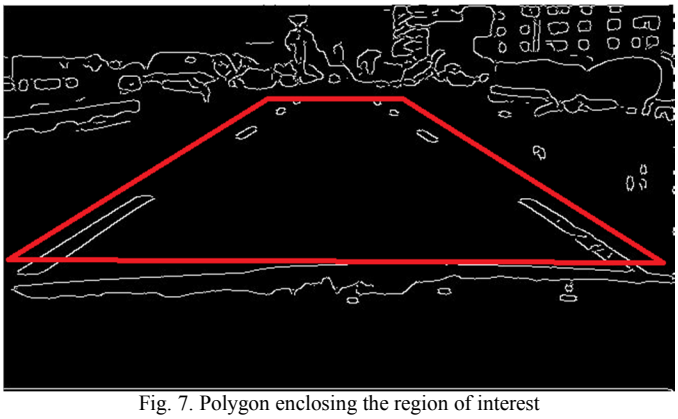


Fig. 7. Polygon enclosing the region of interest

Everything outside the highlighted area is black or null. We can accomplish this by applying a mask which enclosing the lanes just like the red lines do in the above image.

The images with the region of interest mask is shown below:

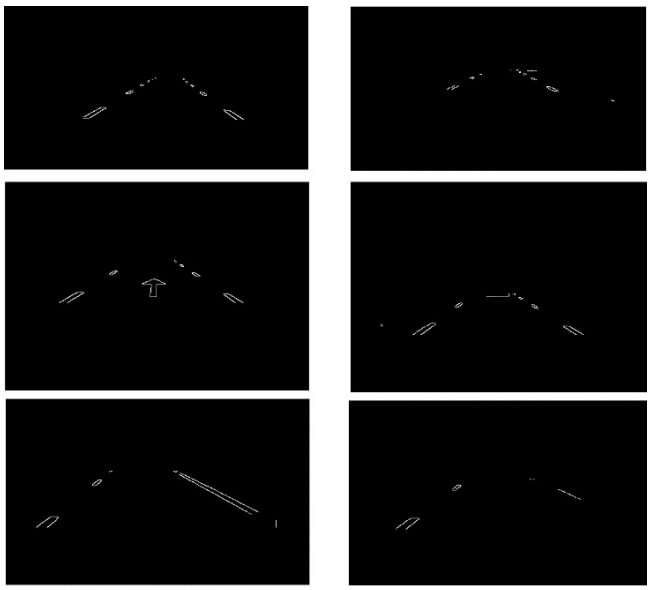


Fig .8. Applying region of interest mask

Now we have lane lines, but we still need the computer to recognize them as lines. Especially, two lines: the left lane and the right lane.

4. Hough Transform Line Detection:

As we can observe from fig 8. the edges of the lanes are actually points or dots so all we need to do is join these dots to create a line and extend the lines till the horizon. The dots can be joined using Hough Transform Line Detection.

The images after applying Hough Transform Line Detection is shown below:

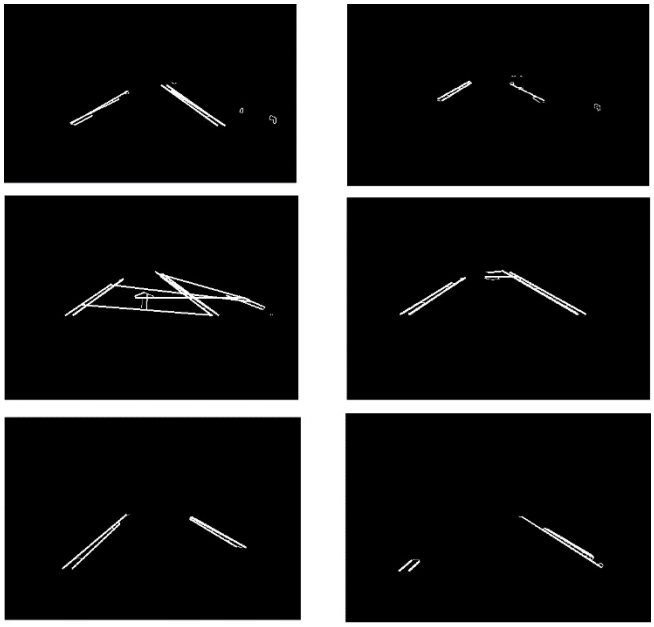


Fig. 9. Applying Hough Transform Line detection

Lastly, we average out the multiple lines detected for a lane lines so that we have only one line for each lane. Next, we extrapolate these lines which basically means extending them till the horizon.

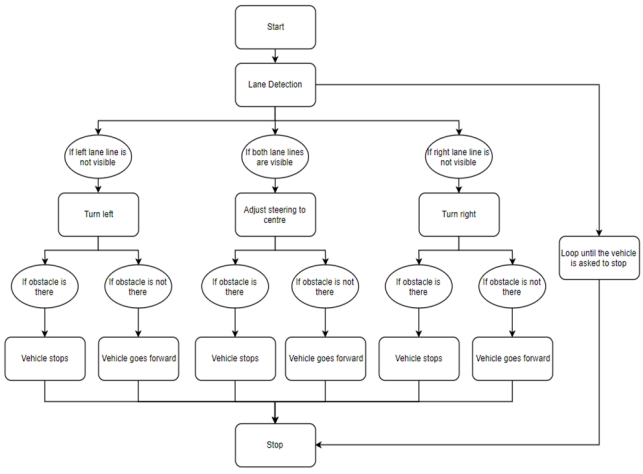


Fig. 10. Lane detection flowchart

The final images with the lane lines being detected are as follows:

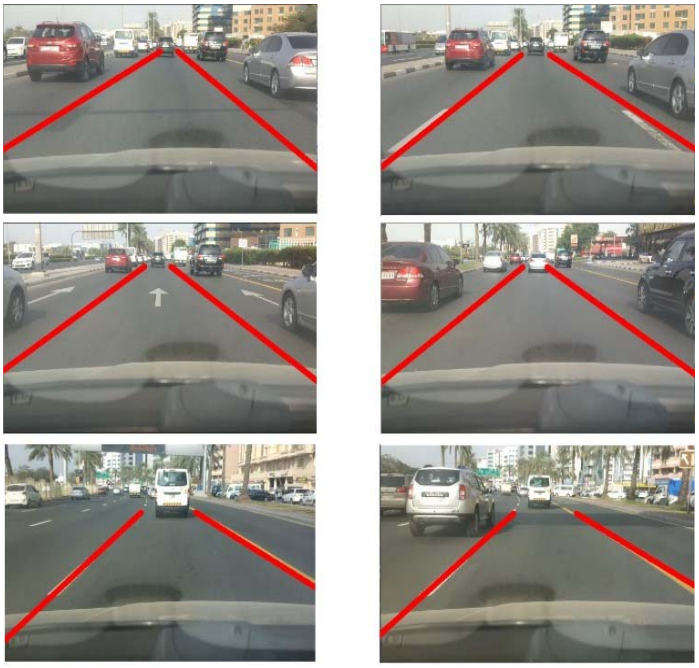


Fig. 11. Extrapolation of Hough transform lines

V. OBJECT DETECTION USING DEEP LEARNING

Object detection is one of the classical issues in computer vision. Recognizing what objects are within a given image and additionally wherever they are within that image may be a challenge. Detection is a lot of complicated downside than classification, which might additionally acknowledge objects however doesn't tell wherever precisely the object is found within the image. YOLO stands for 'You Only Look Once'. It is an object detector that uses a single convolutional neural network for both classification and localization of the object using bounding boxes. YOLO takes a completely different approach [8]. YOLO looks at the image just once but in a clever way as shown below:

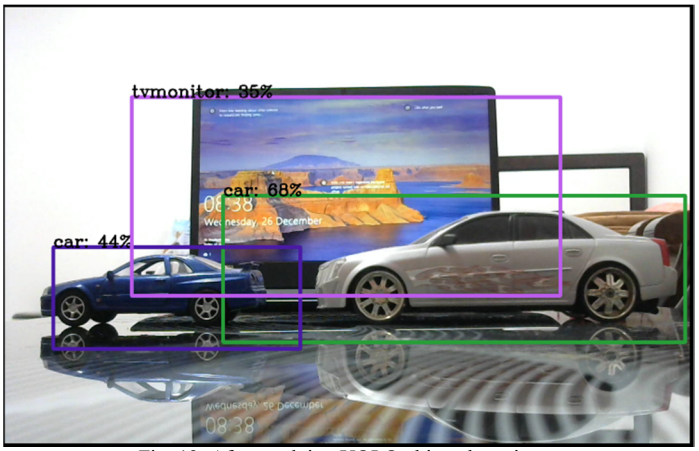


Fig. 12. After applying YOLO object detection

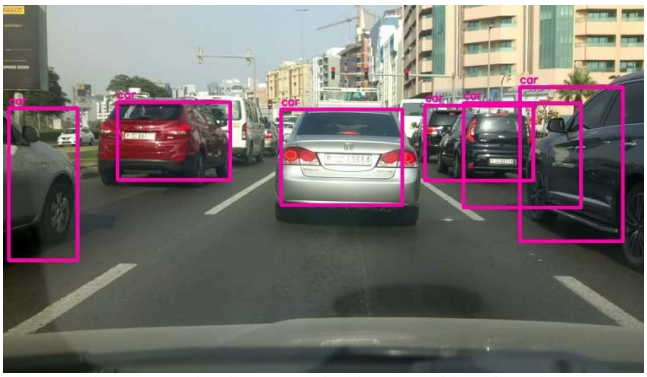


Fig. 13. Another live example of YOLO object detection

VI. Conclusion and Results

Using Image processing and obstacle detection with ultrasonic sensors, we were able to achieve our goal of autonomous mobility vehicle. The image processing part includes lane detection and object recognition to distinguish between pedestrians and objects. The lane detection is the core part of this project since its responsible for making the vehicle maintain its lane. The ultrasonic sensors just detect any obstacle in front and immediately make the vehicle stop within the distance as set by the user. The user would be able to select the speed that he/she would like to travel at since this function is accessible to the user. The user can use the indicators and horn as well whenever required. Using big sensors such as LIDAR etc. this vehicle can be further improvised, and its accuracy can be increased, since LIDAR sensors can map the entire surroundings [9].

We were able to successfully test our autonomous vehicle model in a closed loop environment complete with road tracks. We incorporated the lane detection algorithm and collision avoidance with a high success rate.

We then compiled the lane detection algorithm, collision avoidance system, and also incorporated deep learning algorithms for object detection We can further increase the safety and autonomous level by using a Lidar Sensor which gives a 3-D representation of the vehicles surrounding [9].


Fig. 14. Autonomous Mobility Vehichle

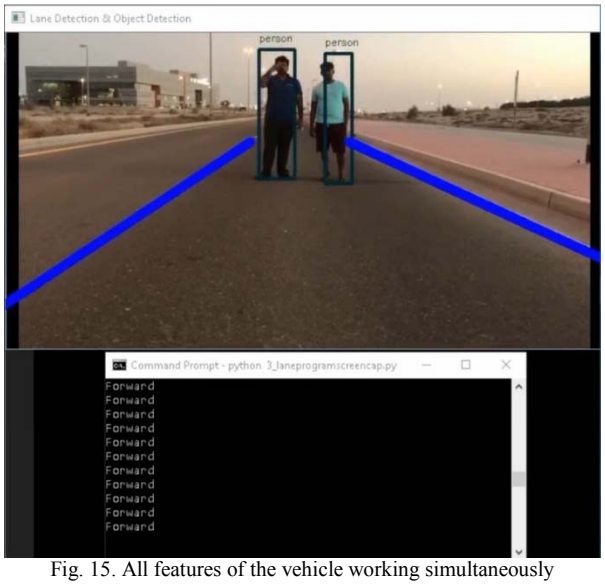


Fig. 15. All features of the vehicle working simultaneously


Fig. 16. Autonomous Mobility vehicle with a passenger

Acknowledgement

The authors would like to acknowledge the staff support of Manipal Academy of Higher Education Dubai for their guidance. The authors would also like to acknowledge the other project members who have played a significant role to make this project a success.

References

[1] Marilyn J. Field and Alan M. Jette, The Future of Disability in America. Washington, DC: National Academic Press, 2007.

[2] Bimbraw, Keshav. (2015). Autonomous Cars: Past, Present and Future - A Review of the Developments in the Last Century, the Present Scenario and the Expected Future of Autonomous Vehicle Technology. ICINCO 2015 - 12th International Conference on Informatics in Control, Automation and Robotics, Proceedings. 1. 191-198. 10.5220/0005540501910198.

[3] A. A. Assidiq, O. O. Khalifa, M. R. Islam and S. Khan, "Real time lane detection for autonomous vehicles," 2008 International Conference on Computer and Communication Engineering, Kuala Lumpur, 2008, pp. 82-88, doi: 10.1109/ICCCE.2008.4580573.

[4] ElectFreaks, "Ultrasonic Ranging Module HC-SR04,"[Online]. Available: http://www.electroschematics.com/8902/hc-sr04-datasheet/.

[5] Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.

[6] Y. Han and E. Oruklu, "Traffic sign recognition based on the NVIDIA Jetson TX1 embedded system using convolutional neural networks," 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), Boston, MA, 2017, pp. 184-187, doi: 10.1109/MWSCAS.2017

[7] I. Culjak, D. Abram, T. Pribanic, H. Dzapo and M. Cifrek, "A brief introduction to OpenCV," *2012 Proceedings of the 35th International Convention MIPRO*, Opatija, 2012, pp. 1725-1730.

[8] Joseph Redmon,Ali Farhadi. “You Only Look Once: Unifies Real Time Object Detection”,2017.

[9] M. Cecotti, H. Kanchwala and N. Aouf, "Autonomous Navigation for Mobility Scooters: a Complete Framework Based on Open-Source Software*," 2019 IEEE Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 2019, pp. 3627-3632, doi: 10.1109/ITSC.2019.8917469.

Estimation of Dynamic Laplacian Eigenvalues in Dynamic Consensus Networks

Fadel Lashhab, Member, IEEE
Department of Electrical Engineering and Computer Science
College of Engineering and Architecture
Howard University, Washington DC, USA.
fadel.lashhab@howard.edu

***Abstract*—In this paper we generalize recent results on networks with static weights by introducing the idea of dynamic networks with real rational weights. Specifically, we consider networks whose nodes are transfer functions (typically integrators) and whose edges are strictly positive real transfer functions representing dynamical systems that couple the nodes. We show that strictly positive realness of the edges is a sufficient condition for dynamic networks to be stable (i.e., to reach consensus). To study the spectral properties of dynamic networks, we introduce the Dynamic Grounded Laplacian matrix, which is used to estimate lower and upper bounds for the real parts of the smallest and largest non-zero eigenvalues of the dynamic Laplacian matrix. These bounds can be used to obtain stability conditions using the Nyquist graphical stability test for undirected dynamic networks controlled using distributed controllers. Numerical simulations are provided to verify the effectiveness of the results.**

***Index Terms*—Dynamic graphs, dynamic consensus networks, dynamic Laplacian, grounded dynamic Laplacian, bounds for the eigenvalues of the dynamic Laplacian matrix, Nyquist graphical stability test, distributed controllers.**

I. Introduction

The idea of consensus in networks has received considerable attention due to a wide array of applications in fields such as robotics, transportation, sensor networks, communication networks, biology, and physics. The focus of this paper is to study a generalization of consensus problems in which the weights of network edges are no longer static gains, but instead are dynamic systems, leading to the notion of *dynamic consensus networks*.

To understand the basic idea of *consensus*, consider the graph or network shown in Fig. 1. This graph consists of three nodes or agents with node variables denoted as v_i, v_j, and v_k and edges e_{ij} connecting the i^{th} and j^{th} nodes. Such graphs can be described as a set of nodes (or vertices) $\mathcal{N} = \{n_i\}$ connected by a set of edges $\mathcal{E} = \{(n_i, n_j) : \ n_i, n_j \in \mathcal{N}\}$. Each edge $\mathcal{E}$ is associated with a flow $f_{ij}(t)$, which in this figure is the same as $f_{ji}(t)$, implied by the bi-directional arrows.

In Fig. 1, we assume the node does some form of processing of its inputs and that edges carry a flow of information or energy from one node to the other. Assuming an external input

978-1-7281-9615-2/20$31.00

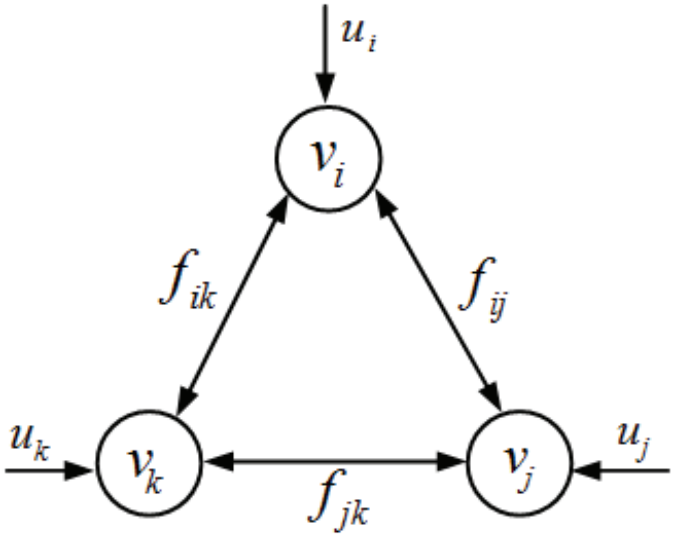


Fig. 1. Graph or network.

u_i to each node, then a typical form of the processing in each node is, for example,

$$\dot{v_i}(t) = u_i(t) - \sum_{j \in \mathcal{N}_i} f_{ij}(t) \tag{1}$$

where $\mathcal{N}_i$ is the set of indices for nodes that communicate with node i. This is called the neighborhood of node i and the nodes in this case are considered to be integrators. Note that a popular form for $f_{ij}(t)$ is

$$f_{ij}(t) = \lambda_{ij}(v_i(t) - v_j(t)), \tag{2}$$

where λ_{ij} is a static gain or interconnection weight between node i and node j. With $u_i(t) = 0$, (1) results in the so-called *consensus protocol*

$$\dot{v_i}(t) = -\sum_{j \in \mathcal{N}_i} \lambda_{ij}(v_i(t) - v_j(t)). \tag{3}$$

The continuous-time linear consensus protocol (3) can be written in matrix form as:

$$\dot{v}(t) = -Lv(t), \tag{4}$$

where $v(t) = [v_1(t), v_2(t), ..., v_n(t)]^T$ and $L = [l_{ij}]$, is the Laplacian matrix, defined by

$$l_{ij} = \begin{cases} \sum_{j \in \mathcal{N}_i} \lambda_{ij} & i = j \\ -\lambda_{ij} & i \neq j \text{ and } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

For the multi-agent consensus problem, the state $v_i \in \mathbb{R}^1$ represents the belief of agent i about a so-called global consensus variable v. These beliefs are updated using communications with their neighbors according to the consensus protocol (3). A key result is that the solution of $\dot{v}(t) = -Lv(t)$ gives $v_i \to v^*$, where v^* is a constant, if the static graph is connected [24]. This specific fact has been the basis of much of the literature related to consensus problems. The focus of this paper is to study a generalization of consensus problems whereby the weights of the network edges are no longer modeled as static gains, but instead are represented as dynamic systems coupling the nodes.

The consensus protocol given in (3) and its variants have been studied extensively in the literature and have been applied in a number of areas, notably to formation motion control when the agents are mobile. We have been motivated to consider this extension to transfer functions weights from several applications. In particular, the thermal model of heat flows in buildings can be shown to admit a generalized dynamic consensus network over a directed graph [22]. In another example, modeling the load frequency control (LFC) network of an electrical power grid can be viewed as a single-integrator consensus network with dynamic interconnection coefficients $\lambda_{ij} = G_i(s)a_{ij}$ [16]. To further motivate the concept from an application, in the next section we show how an electrical circuit whose nodes are grounded capacitors can be modeled as an undirected dynamic consensus network.

This paper offers several distinct contributions. First, we show that the nodes variables in dynamic consensus networks whose nodes are integrators and whose edges are strictly positive real transfer functions will converge to a common consensus value. Second, we develop methods for estimating the bounds of the eigenvalues of a dynamic Laplacian matrix. This is accomplished by introducing the idea of a grounded dynamic Laplacian sub-matrix as a means of reducing the order of complexity of computation that would be necessary in the case of the full dynamic Laplacian matrix. Third, we use the estimated bounds of the eigenvalues of a dynamic Laplacian to analyze the consensus and stability of a dynamic consensus network under distributed feedback control.

The paper is organized as follows: Section 2 presents a motivating physical example of a dynamic consensus network. In Section 3 we investigate the consensus conditions of dynamic consensus networks with integrator nodes and strictly positive real (SPR) edges. For such graphs, we can introduce what we call the dynamic Laplacian and dynamic Grounded Laplacian of the dynamic graphs. In Section 4, lower and upper bounds for the largest and smallest non-zero eigenvalues of the dynamic Laplacian matrix are derived. Finally, using these bounds, we characterize in Section 5 the stability of a closed loop system consisting of an undirected dynamic consensus network controlled by distributed controllers, and a numerical example is provided to illustrate the results.

II. Dynamic Consensus Network

A generalization of static consensus networks can be motivated when the consensus network consists of an interconnection of physical systems (e.g., an electrical network that requires a common operating condition such as voltage or frequency) or in cases when the communication between agents in a consensus network includes dynamic behavior (e.g., delays). This paper studies a generalization of the static consensus problem whereby the weights are no longer positive gains, but instead represent strictly-positive real dynamic systems. The use of positive real (PR) transfer functions is motivated by the study of dynamic networks that represent linear electrical circuits composed of resistors, inductors and capacitors. A strictly positive real transfer (SPR) function corresponds to the driving input impedance in dissipative networks that contain resistors, inductors and capacitors [28].

Consider a dynamic network whose node equations can be written in the time domain for the node i, $i = 1, 2, ..., n$, as

$$\dot{v}_i(t) = -\sum_{j \in \mathcal{N}_i} y_{ij}(t) * [v_i(t) - v_j(t)], \quad (6)$$

where v_i is a consensus variable of the node i, $y_{ij}(t)$ is the impulse response of the edge e_{ij}, and "$*$" denotes the convolution operation. Taking the Laplace transform of (6) results in

$$V_i(s) = -\frac{1}{s}\sum_{j \in \mathcal{N}_i} Y_{ij}(s)[V_i(s) - V_j(s)] + \frac{1}{s}v_i(0), \quad (7)$$

where $Y_{ij}(s)$ is the Laplace transform of $y_{ij}(t)$. We refer to (6) or (7)as a *dynamic consensus* protocol. Such a transfer function can be viewed as a weight on the edge connecting node j to node i. As a special case, we consider graphs where these weights $Y_{ij}(s)$ are strictly-positive real transfer functions.

We now consider graphs formed by nodes with multiple terminals that are connected by weighted edges where the weights are represented as transfer functions taken from real rational functions analytic in the right half of the complex plane. These graphs are connected with the dynamic consensus protocol introduced earlier, and the dynamic Laplacian matrix is introduced. To illustrate how such a network can arise in applications, consider a simple electrical network shown in Fig. 2. This circuit can be represented as a dynamic graph by viewing the capacitors as nodes of the dynamic graph and the connections between capacitors as edges in which each edge is assigned a real rational transfer function $Z_{ij}(s)$ as its impedance. The node variables (v_i, $\forall i = 1, ..., N$) are voltages across the capacitors, and current is the flow along the edges between nodes. The weight of an edge can be viewed as the admittance of that edge, which we will denote $Y_{ij}(s) = \frac{1}{Z_{ij}(s)}$ for a given edge e_{ij}. If there is a potential difference $(v_i - v_j)$

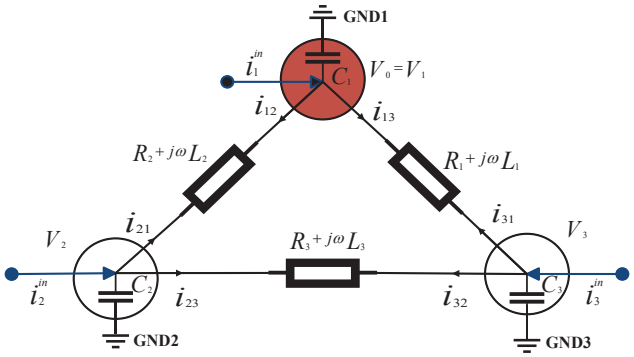


Fig. 2. Electrical network as dynamic graph.

between the nodes i and j, an electrical current i_{ij} will flow across the edge e_{ij} from i to j according to Ohm's law, which states that the potential drop across an edge e_{ij} must be equal to the product (in the complex frequency domain) of its impedance and the current flowing through it:

$$I_{ij}(s) = Y_{ij}(s)[V_i(s) - V_j(s)], \tag{8}$$

where $I_{ij}(s)$ and $V_i(s)$ are the Laplace transform of the current $i_{ij}(t)$ and the node voltage $v_i(t)$, respectively. Using Kirchhoff's law, it can be established that the algebraic sum of all of the currents entering a node i is equal to the algebraic sum of all currents leaving that node. Therefore, for Fig. 2, the dynamic model of each node is obtained as:

$$\begin{aligned} C\dot{v}_1(t) &= i_1^{in}(t) - (i_{12}(t) + i_{13}(t)); \\ C\dot{v}_2(t) &= i_2^{in}(t) - (i_{21}(t) + i_{23}(t)); \\ C\dot{v}_3(t) &= i_3^{in}(t) - (i_{31}(t) + i_{32}(t)). \end{aligned} \tag{9}$$

Noting that

$$i_{ij}(t) = -i_{ji}(t), \tag{10}$$

and, by combining (8) with the Laplace transformed version of (9) and setting $C = 1F$, we obtain:

$$\begin{aligned} V_i(s) =& \frac{1}{s} I_i^{in}(s) - \frac{1}{s} \sum_{j \in \mathcal{N}_i} Y_{ij}(s)[V_i(s) - V_j(s)] \\ &+ \frac{1}{s} v_i(0), \forall i = 1, 2, 3, \end{aligned} \tag{11}$$

where $\mathcal{N}_i$ is the set of neighbors node i. For the autonomous case ($I_i^{in}(s) = 0$), the dynamic model of each node with zero initial conditions ($v_i(0) = 0$) can then be written as:

$$V_i(s) = -\frac{1}{s} \sum_{j \in \mathcal{N}_i} Y_{ij}(s)[V_i(s) - V_j(s)], \forall i = 1, 2, 3. \tag{12}$$

Defining the node-potential vector $V(s) = [V_1(s), V_2(s), ..., V_N(s)]^T$, the overall system can be represented in the frequency domain by:

$$V(s) = -\frac{1}{s} L(s) V(s). \tag{13}$$

In this expression the matrix $L(s) = [l_{ij}(s)]$ is given as:

$$l_{ij}(s) = \begin{cases} \sum_{j \in \mathcal{N}_i} Y_{ij}(s) & i = j \\ -Y_{ij}(s) & i \neq j \text{ and } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

This leads to:

$$L(s) = \begin{bmatrix} Y_{12}(s) + Y_{13}(s) & -Y_{12}(s) & -Y_{13}(s) \\ -Y_{12}(s) & Y_{12}(s) + Y_{23}(s) & -Y_{23}(s) \\ -Y_{13}(s) & -Y_{23}(s) & Y_{13}(s) + Y_{23}(s) \end{bmatrix}. \tag{15}$$

$L(s)$ is defined in this way as a dynamic Laplacian matrix of the dynamic graph G. In the sequel, we will also refer to the restriction $L(j\omega) = L(s)$ where $s = (0 + j\omega)$.

III. Consensus Conditions of Networks with Integrator Nodes and Strictly Positive Real (SPR) Edges

In this section we show that the strictly positive realness of the edges is a sufficient condition for dynamic networks of the form (12) to reach a consensus (i.e., all node variables converge to the same constant value in the steady-state).

In this paper, the edges are assumed to be strictly positive real. Thus, for the dynamic network (12) with strictly-positive real edges, we must look to the strictly positive realness and positive definiteness of the complex matrix. To proceed, we first define the concepts of strictly-positive real edges transfer functions and positive definiteness for complex matrices, and we give two definitions and two lemmas needed for our result:

Definition 3.1: [13] An $n \times n$ complex matrix A is called positive definite (PD) matrix if $Re[x^H Ax] > 0$ and positive semidefinite (PSD) if $Re[x^H Ax] \geq 0$, for all complex vector $x \in \mathbb{C}^n$, where x^H denotes the conjugate transpose of the vector x.

Definition 3.2: [28] An $n \times n$ proper rational transfer function matrix $G(s)$ is positive real if

1) Poles of all elements of $G(s)$ are in Re$[s] \leq 0$.
2) For all real ω for which $j\omega$ is not a pole of any element of $G(s)$, the matrix $G(j\omega) + G^T(-j\omega)$ is positive semidefinite (PSD).
3) Any purely imaginary pole $j\omega$ of any element of $G(s)$ is a simple pole and the residue matrix $\lim_{s \to j\omega}(s - j\omega)G(s)$ is positive semidefinite Hermitian.

In addition, a real rational transfer function $h(s)$ is said to be strictly positive real (SPR) if there exists ϵ_0 such that $h(s-\epsilon)$ is positive real for $0 < \epsilon < \epsilon_0$.

Lemma 3.3: [13] A necessary and sufficient condition for a complex matrix A to be PD or PSD is that the Hermitian part $H(A) = \frac{1}{2}(A + A^H)$ must be PD or PSD, respectively.

Lemma 3.4: [13] Let $A \in \mathbb{C}^{n \times n}$. If $A + A^H$ is PD or PSD, then all the eigenvalues of matrix A have positive or non-negative real parts, respectively.

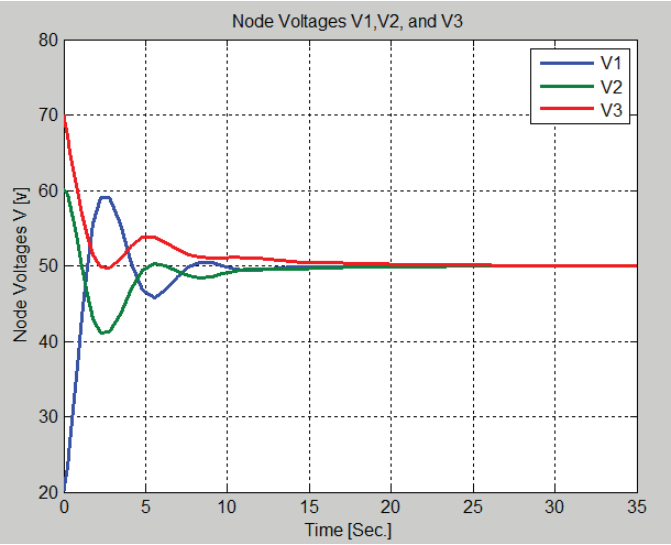


Fig. 3. Simulation of the node voltages of the dynamic network.

Using these definitions and lemmas, for the proposed dynamic network (12) we can give the following result:

Theorem 3.5: [14] Given system (13) for the dynamic consensus protocol (12) with edges that are SPR transfer functions, the following statements are true:

1) The dynamic Laplacian $L(s)$ is a complex, symmetric, positive semidefinite (CSPSD) matrix.
2) The real parts of the eigenvalues of $L(j\omega)$ are non-negative ($\text{Re}[\lambda_i(L(j\omega))] \geq 0, \forall i \in 1,\ldots,n$), with $0 = \lambda_1(L(j\omega)) \leq \text{Re}[\lambda_2(L(j\omega))] \leq \ldots \leq \text{Re}[\lambda_n(L(j\omega))]$.
3) If $L(0)$ (the dynamic Laplacian $L(s)$ evaluated at $s=0$) is the Laplacian matrix of a connected graph, then $s=0$ is a distinct solution of $\det[sI_N + L(s)] = 0$ and all node variables will reach consensus under the protocol (12).

As an example, consider the dynamic network of Fig. 2 with SPR transfer functions $Y_{ij}(s) = \frac{1}{Z_{ij}(s)}$ assigned for the edges, and for simplicity consider $Z_{ij} = R + j\omega L, R = 1\Omega$, and $L = 1H$. The dynamic Laplacian $L(s)$ can be written as

$$L(s) = \begin{bmatrix} \frac{2}{1+s} & -\frac{1}{1+s} & -\frac{1}{1+s} \\ -\frac{1}{1+s} & \frac{2}{1+s} & -\frac{1}{1+s} \\ -\frac{1}{1+s} & -\frac{1}{1+s} & \frac{2}{1+s} \end{bmatrix} \Rightarrow L(0) = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.$$

From the above, it is readily verified that $-L(0)$ is the Laplacian matrix of a connected graph. In fact, $L(0)$ is a connected interconnection matrix (see Definition [22]). Applying Theorem 3.5, we expect that all node variables converge to consensus. This is verified in Fig. 3, which shows a simple simulation of the dynamic consensus protocol (12), with initial conditions for the node potentials taken to be $V(0) = [20, 60, 70]^T$.

IV. Bounding the Eigenvalues of the Dynamic Laplacian Matrix

In this section, we develop a method to estimate the bounds of the eigenvalues of the dynamic Laplacian matrix $L(j\omega)$ for all ω. Our approach uses the idea of the *Dynamic Grounded Laplacian*. Motivated by the static case, we introduce the *Dynamic Grounded Laplacian* matrix to estimate lower and upper bounds for the real parts of the smallest and largest non-zero eigenvalues of the dynamic Laplacian matrix. We begin by defining the Dynamic Grounded Laplacian matrix. We then show how to obtain the lower and upper bounds for the eigenvalues of the dynamic Laplacian matrix.

A. Dynamic Grounded Laplacian Matrix

The Dynamic Grounded Laplacian matrix can be defined from the dynamic Laplacian matrix by extending the definition of the Grounded Laplacian matrix in the case of static networks (e.g., see [2] and [3]) to the case of dynamic networks.

To illustrate the idea, we will define the Dynamic Grounded Laplacian matrix for the undirected dynamic graph that represents a simple electrical network as shown in Fig. 2. As we have mentioned earlier, this network can be modeled as a dynamic graph by representing the capacitors on the vertices as integrator nodes and the transmission lines or the branches between the nodes as SPR transfer functions edges.

Using Kirchhoff's law, it can be established that the dynamic Laplacian matrix $L(j\omega)$ expresses a mapping from input currents to node potentials as

$$\begin{bmatrix} I_1^{in}(j\omega) \\ I_2^{in}(j\omega) \\ I_3^{in}(j\omega) \end{bmatrix} = L(j\omega) \begin{bmatrix} V_1(j\omega) \\ V_2(j\omega) \\ V_3(j\omega) \end{bmatrix}, \tag{16}$$

where $L(j\omega)$ is defined by (15), where $Y_{ij}(j\omega)$ is the admittance between nodes i and j.

For a given dynamic network, assume that the node V_1 is connected to the ground (this node is colored red in the electrical network shown in Fig. 2). Grounding node V_1 means that the potential and the current of this node are zeros ($V_1(j\omega) = 0$ and $I_1^{in}(j\omega) = 0$).

Substituting $V_1(j\omega) = 0$ and $I_1^{in}(j\omega) = 0$ in the equation (16), we get

$$\begin{bmatrix} I_2^{in}(j\omega) \\ I_3^{in}(j\omega) \end{bmatrix} = \underbrace{\begin{bmatrix} Y_{12}(j\omega) + Y_{23}(j\omega) & -Y_{23}(j\omega) \\ -Y_{23}(j\omega) & Y_{13}(j\omega) + Y_{23}(j\omega) \end{bmatrix}}_{L_0(j\omega)} \begin{bmatrix} V_2(j\omega) \\ V_3(j\omega) \end{bmatrix}. \tag{17}$$

We call $L_0(j\omega)$, defined in this way, a Dynamic Grounded Laplacian matrix. We give this matrix this name because it is defined from the dynamic Laplacian by grounding the subset nodes $\mathbf{V_0} = V_1$ in this case.

These ideas can be generalized: Given a dynamic graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, where $\mathbf{V}$ is a set of n nodes; $\mathbf{E} \subset \mathbf{V} \times \mathbf{V}$ is a set of m edges, and given a subset of grounded nodes $\mathbf{V_0} \subset \mathbf{V}$ consisting of $n_g < n$ nodes, $n_0 = n - n_g$, then the *Dynamic Grounded Laplacian* matrix $L_0(j\omega) \in \mathbb{C}^{n_0 \times n_0}$ is defined from the dynamic Laplacian matrix $L(j\omega)$ by removing all rows and columns corresponding to the nodes in $\mathbf{V_0}$. Note that this must be done in such a way that $L_0(j\omega)$ corresponds to a connected

graph. The following lemma describes the properties of the dynamic Grounded Laplacian $L_0(j\omega)$.

Lemma 4.1: For an undirected, connected dynamic graph with SPR edges, the Dynamic Grounded Laplacian matrix $L_0(j\omega) \in \mathbb{C}^{n_0 \times n_0}$ is a complex, symmetric, positive definite (CSPD) matrix that is always invertible for all ω.

Proof:

Since the dynamic graph is undirected, the dynamic Laplacian matrix is symmetric $Y_{ij}(j\omega) = Y_{ji}(j\omega)$ and then $L_0(j\omega)$ is a complex symmetric (CS) matrix.

Now, we can show $L_0(j\omega)$ is PD by examining the positive definiteness of the Hermitian part of $L_0(j\omega)$ (see Definition 3.1),

$$H(L_0(j\omega)) = \frac{1}{2}(L_0(j\omega) + L_0(j\omega)^H) = \frac{1}{2}(L_0(j\omega) + \bar{L}_0(j\omega)) = \text{Re}[L_0(j\omega)].$$

Based on Lemma 3.3, the positive definiteness of $L_0(j\omega)$ is due to $x^T \text{Re}[L_0(j\omega)]x = \sum_{j \in \mathcal{N}_i^0} \text{Re}[Y_{ij}(j\omega)]|x_i - x_j|^2 > 0$ for all $x \in \mathbb{R}^{n_0}$, where $\mathcal{N}_i^0$ is the set of neighborhood nodes to the node i in the remaining graph after grounding. Thus, the Dynamic Grounded Laplacian matrix $L_0(j\omega)$ is always invertible for all ω. ■

B. Finding Bounds for the Eigenvalues of the Dynamic Laplacian Matrix

In this subsection, the bounds for the eigenvalues of the dynamic Laplacian are estimated using the properties of the dynamic Laplacian and Dynamic Grounded Laplacian matrices.

1) Lower Bound for the Smallest Non-Zero Eigenvalue of the Dynamic Laplacian Matrix: The following theorem provides a lower bound of the dynamic Laplacian matrix. Later this bound will be used to characterize the stability of the closed loop system in the case of distributed controllers.

Theorem 4.2: Let $L(j\omega) \in \mathbb{C}^{n \times n}$ be the dynamic Laplacian matrix for an undirected dynamic graph with SPR edges. Let $\lambda_i(L(j\omega))$ and $\lambda_i(L_0(j\omega))$ denote the eigenvalues of $L(j\omega)$ and $L_0(j\omega)$, respectively, for any choice of a boundary set of grounded nodes **V0**, for which the graph is connected. Assume that $\text{Re}[\lambda_1(L(j\omega))] \leq \text{Re}[\lambda_2(L(j\omega))] \leq ... \leq \text{Re}[\lambda_n(L(j\omega))]$ and $\text{Re}[\lambda_1(L_0(j\omega))] \leq \text{Re}[\lambda_2(L_0(j\omega))] \leq ... \leq \text{Re}[\lambda_{n_0}(L_0(j\omega))]$, $n_0 < n$. Then a lower bound for the real part of the smallest non-zero eigenvalue of the dynamic Laplacian is given by

$$\min_{\omega} \text{Re}[\lambda_2(L(j\omega))] \geq \frac{1}{\|\text{trace}((\text{Re}[L_0(j\omega)])^{-1})\|_\infty}. \quad (18)$$

Proof:

The proof of this Theorem is given in the following steps:

1) We begin the proof by finding lower bounds for the eigenvalues of the Hermitian part of a Dynamic Grounded Laplacian matrix $\lambda_i(H(L_0(j\omega)))$:
Based on Lemma A.1, these bounds can be obtained as
$$\lambda_i(H(L_0(j\omega))) \geq \frac{1}{\text{trace}((\text{Re}[L_0(j\omega)])^{-1})}, \quad (19)$$
for all $\omega, i \in 1, 2, ..., n$.
2) In this step, we derive the inequality that relates the eigenvalues of the matrices $H(L_0(j\omega))$ and $H(L(j\omega))$: Applying the interlacing eigenvalue Theorem A.2, we can conclude that the eigenvalues of $H(L(j\omega)) \in \mathbb{R}^{n \times n}$ and $H(L_0(j\omega)) \in \mathbb{R}^{n_0 \times n_0}$ are interlaced for $i = 1, 2, ..., n_0$, as
$$\lambda_i(H(L(j\omega))) \leq \lambda_i(H(L_0(j\omega))) \leq \lambda_{i+(n-n_0)}(H(L(j\omega))). \quad (20)$$
Combining (19) and (20), we get
$$\lambda_i(H(L(j\omega))) \geq \frac{1}{\text{trace}((\text{Re}[L_0(j\omega)])^{-1})}, \quad (21)$$
for all ω and $i = (n - n_0) + 1, ..., n$.
3) Here, we derive the relationship between the real part of the smallest and largest nonzero eigenvalues of $L(j\omega)$ and $H(L(j\omega))$. This relationship can be achieved using Theorem A.3 as
$$\sum_{i=n-k}^{n} \text{Re}[\lambda_i(L(j\omega))] \leq \sum_{i=n-k}^{n} \lambda_i(H(L(j\omega))), \quad (22)$$
$k = 0, 1, ..., n-1$, with equality for $k = n - 1$.
Based on Mirsky Theorem A.3, the relationship between the eigenvalues of the matrices $L(j\omega)$ and $H(L(j\omega))$ can be obtained by substituting $k = n - 3$, $k = n - 1$, and $k = 0$ in (22) and then combining the result. Thus, the inequalities that characterize the relationship between the eigenvalues of these matrices are given by Lemma A.4 as
$$\text{Re}[\lambda_2(L(j\omega))] \geq \lambda_2(H(L(j\omega))) \quad (23)$$
and
$$\text{Re}[\lambda_n(L(j\omega))] \leq \lambda_n(H(L(j\omega))), \quad (24)$$
where $\text{Re}[\lambda_2(L(j\omega))]$ and $\text{Re}[\lambda_n(L(j\omega))]$ are the real part of the smallest and largest nonzero eigenvalues of $L(j\omega)$ and $H(L(j\omega))$, respectively. Note that these result are obtained for the case of a connected dynamic graph with SPR edges, which is mean $\text{Re}[\lambda_1(L(j\omega))] = \lambda_1(H(L(j\omega))) = 0$.
4) Finally, a lower bound for the real part of the smallest non-zero eigenvalue of $L(j\omega)$ can be obtained by combining (21) and (23) as
$$\text{Re}[\lambda_2(L(j\omega))] \geq \frac{1}{\text{trace}((\text{Re}[L_0(j\omega)])^{-1})}, \forall\omega. \quad (25)$$
Taking the minimum values of both sides in the above inequality over all ω, yields
$$\min_{\omega} \text{Re}[\lambda_2(L(j\omega))] \geq \min_{\omega}[\frac{1}{\text{trace}((\text{Re}[L_0(j\omega)])^{-1})}];$$
$$\min_{\omega} \text{Re}[\lambda_2(L(j\omega))] \geq \frac{1}{\max_{\omega}[\text{trace}((\text{Re}[L_0(j\omega)])^{-1})]}. \quad (26)$$

Therefore, a lower bound for the real part of the smallest non-zero eigenvalue of the dynamic Laplacian matrix is given by

$$\min_{\omega} \mathrm{Re}[\lambda_2(L(j\omega))] \geq \frac{1}{\|\mathrm{trace}((\mathrm{Re}[L_0(j\omega)])^{-1})\|_\infty}. \quad (27)$$

■

2) An Upper Bound for the Largest Eigenvalue of the Dynamic Laplacian Matrix: The following theorem provides an upper bound for the real part of the largest eigenvalue of the dynamic Laplacian matrix:

Theorem 4.3: Let $L(j\omega) \in \mathbb{C}^{n\times n}$ be the dynamic Laplacian matrix for an undirected dynamic graph with SPR edges. Let $\lambda_i(L(j\omega))$ denote the eigenvalues of $L(j\omega)$, ordered so that $\mathrm{Re}[\lambda_1(L(j\omega))] \leq \mathrm{Re}[\lambda_2(L(j\omega))] \leq \ldots \leq \mathrm{Re}[\lambda_n(L(j\omega))]$. Then an upper bound for the real part of the largest eigenvalue of the dynamic Laplacian matrix is given by

$$\max_{\omega} \mathrm{Re}[\lambda_n(L(j\omega))] \leq 2 \max_i \|\mathrm{Re}[D(j\omega)(i,i)]\|_\infty, \quad (28)$$

where $D(j\omega)(i,i)$ is the ith diagonal element of the dynamic Degree matrix of the dynamic graph.

Proof:

First, we show that the $H[2D(j\omega)-L(j\omega)]$ is a PSD matrix for all ω. Because of the positivity realness of the edges and the symmetry property (because the assumed dynamic graph uses undirected topology), we can write

$$\begin{aligned} &H[2D(j\omega) - L(j\omega)] = H[2D(j\omega) - (D(j\omega) - A(j\omega)] \\ &= H[D(j\omega) + A(j\omega)] = \mathrm{Re}[D(j\omega) + A(j\omega)], \end{aligned} \quad (29)$$

where $D(j\omega)$ and $A(j\omega)$ are the dynamic degree and adjacency matrix of the dynamic graph. In addition, the symmetry property of the matrix $[D(j\omega)+A(j\omega)]$ can be easily verified from the following definition: $[D(j\omega) + A(j\omega)] = [l_{ij}(j\omega)]$, where

$$l_{ij}(j\omega) = \begin{cases} \sum_{j\in\mathcal{N}_i} Y_{ij}(j\omega) & i = j \\ +Y_{ij}(j\omega) & i \neq j \text{ and } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise.} \end{cases} \quad (30)$$

Note that the matrix defined above is different from the dynamic Laplacian matrix $L(s)$. The difference is the off-diagonal terms are positive.

By Lemma 3.3, the necessary and sufficient condition for the matrix $H(2D(j\omega) - L(j\omega)) = \mathrm{Re}[D(j\omega) + A(j\omega)]$ to be a PSD matrix is that the real part of $[D(j\omega) + A(j\omega)]$ be a PSD matrix.

Based on Definition 3.1 and (30), we can write

$$x^T \mathrm{Re}[D(j\omega) + A(j\omega)]x = \sum_{j\in\mathcal{N}_i} \mathrm{Re}[Y_{ji}(j\omega)](x_i + x_j)^2 \geq 0,$$

for all vectors $x \in \mathbb{R}^n$. Because of the strictly positive realness of the edges, then $\mathrm{Re}[Y_{ji}(j\omega)]$ is non-negative $\forall\omega \in \mathbb{R}$. Thus, by Definition 3.1, $\mathrm{Re}[D(j\omega)+A(j\omega)]$ is a PSD matrix. Therefore, the matrix $H[2D(j\omega) - L(j\omega)]$ is also a PSD matrix. Hence,

$$H[2D(j\omega) - L(j\omega)] \geq 0 \Leftrightarrow H(L(j\omega)) \leq 2H(D(j\omega)). \quad (31)$$

From the above inequality, we can write

$$\lambda_i(H(L(j\omega))) \leq 2\lambda_i(H(D(j\omega))) = 2\mathrm{Re}[D(j\omega)(i,i)]. \quad (32)$$

Thus, the largest eigenvalue of the Hermitian part of the dynamic Laplacian can be upper-bounded by

$$\begin{aligned} &\lambda_n(H(L(j\omega))) \leq 2\max_i(\max_\omega \mathrm{Re}[D(j\omega)(i,i)]) \\ \Leftrightarrow \quad &\lambda_n(H(L(j\omega))) \leq 2\max_i \|\mathrm{Re}[D(j\omega)(i,i)]\|_\infty. \end{aligned} \quad (33)$$

Combining (24) and (33), we get

$$\mathrm{Re}[\lambda_n(L(j\omega))] \leq 2\max_i \|\mathrm{Re}[D(j\omega)(i,i)]\|_\infty. \quad (34)$$

Thus,

$$\max_{\omega} \mathrm{Re}[\lambda_n(L(j\omega))] \leq 2\max_i \|\mathrm{Re}[D(j\omega)(i,i)]\|_\infty.$$

■

The lower and upper bounds for the eigenvalues of the dynamic Laplacian matrix are useful for stability analysis of dynamic consensus networks under feedback control. In the next section we show how these bound can be used to establish the stability of an undirected dynamic consensus network under distributed control.

V. Stability Analysis of Dynamic Consensus Networks with Distributed Controllers

In this section, we provide conditions for the stability of undirected dynamic networks controlled using distributed controllers. The bounds of the eigenvalues of the dynamic Laplacian matrix $\lambda_i(L(j\omega))$, which can be found using the methodology from the previous section, can be used to analyze the stability of such dynamic networks using the Nyquist graphical-stability test [12]. The authors of [7] and [12] applied tools from graph theory to relate the topology of a static network to the stability of the overall system. They also provided a Nyquist criterion, which uses the eigenvalues of the static Laplacian matrix to determine the effect of the static graph on stability. Here, we extend this stability criterion from static networks to dynamic networks. Before analyzing the stability of the dynamic consensus networks, we will formulate our problem for the proposed dynamic network (undirected dynamic consensus network with identical LTI nodes and dynamic edges).

A. Problem Formulation

Consider a dynamic network with N identical nodes and with dynamic edges e_{ij} (here, we assume the dynamics of each edge is a SPR transfer function), as in Fig. 2, where the voltage of the ith node is represented as $v_i, i = 1, \ldots, N$. Furthermore, we assume that the nodes have identical dynamics $p(s)$ and controllers $k(s)$.

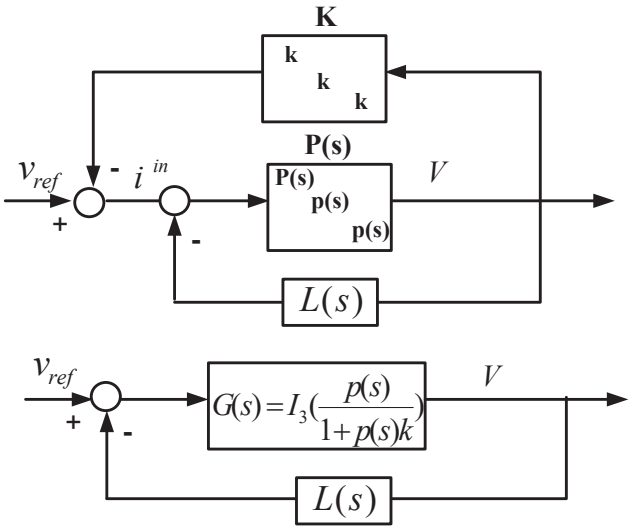


Fig. 4. Upper figure: Block diagram of the controlled system. Lower figure: Equivalent standard feedback diagram.

Consider the dynamic network Fig. 2 as defined in (11) and let the integrator nodes be replaced by a common transfer function such as $p_i(s) = p(s)$. For example, assume that the integrator nodes can be replaced by filters with a common transfer function $p_i(s) = p(s) = \frac{\beta}{s+\alpha}$. Thus, the dynamics of each node can be rewritten in the time domain as

$$\frac{1}{\beta}(\dot{v}_i(t)+\alpha v_i(t)) = i_i^{in}(t)-\sum_{j\in\mathcal{N}_i} y_{ij}(t)*(v_i(t)-v_j(t)), \quad (35)$$

$\forall i = 1,\dots,N$. Now suppose the node voltage v_i of each node is controlled by a static controller k, that determines the input $i_i^{in}(t)$. Then the overall system can be expressed in the frequency domain as

$$\begin{aligned} P(s)^{-1}V(s) &= I^{in}(s) - L(s)V(s) \\ I^{in}(s) &= v^{ref} - KV(s), \end{aligned} \quad (36)$$

where $P(s) = \text{diag}(p(s), p(s), \dots p(s))$, $K = \text{diag}(k(s), k(s), \dots k(s))$, where $k(s) = k$ and $V(s)$, $I^{in}(s)$ are the Laplace transform of the vectors $v(t)$ and $i^{in}(t)$, respectively. The control scheme of this problem can be depicted as shown in the upper part of Fig. 4 (for clarity, shown for $N = 3$). We will apply the Nyquist criterion for this block diagram for our controller design, using the estimated bounds of the eigenvalues of the dynamic Laplacian matrix. Thus, the upper block diagram is first transformed into an equivalent standard feedback diagram as shown in the part of Fig. 4. Note that the identical controller $k(s) = k$ in the equivalent standard feedback diagram is located in the forward path ($G(s)$).

The closed-loop transfer function $H(s)$ from $v^{ref}(s)$ to $V(s)$ can be expressed as:

$$H(s) = G(s)(I_3 + G(s)L(s))^{-1}, \quad (37)$$

where $G(s) = I_N \frac{p(s)}{1+p(s)k}$.

The feedback matrix ($L(s)$) in the equivalent standard feedback diagram of the controlled dynamic network in the lower part of Fig. 4 is a dynamic matrix, whereas this matrix is a constant in the static case. This leads us to proceed with bounding the eigenvalues of the dynamic Laplacian matrix in order to locate them for analyzing the stability of the closed loop system using the Nyquist criterion. These bounds can be predicted for static networks based on properties of the graph [7]. The authors of [7]] considered locations of the eigenvalues of the Laplacian matrix to investigate the stability of a consensus network. The region of reciprocal eigenvalues of the Laplacian matrix can be defined using the estimated regions of the smallest and largest real parts of the non-zero eigenvalues of the dynamic Laplacian matrix as introduced in the previous section. Using these estimated regions, the stability conditions of the closed loop system $H(j\omega)$ can be analyzed with respect to the eigenvalues of $L(j\omega)$ ($\lambda_i(L(j\omega))$) as we will see in the next section.

B. Stability Analysis

Here we analyze the stability of the closed-loop system $H(s)$ with respect to the eigenvalues of $L(j\omega)$ ($\lambda_i(L(j\omega))$). Using Nyquist's graphical-stability test, the stability of the closed loop system (37) with identical node dynamics $p(s)$ and controller $k(s) = k$ can be analyzed as follows:

By Theorem 2 in [10], it can be seen that closed loop system $H(j\omega)$ in (37) is stable if and only if the Nyquist plot of

$$\prod_{i=1}^{N}[1 + \lambda_i(L(j\omega))G(j\omega)] \quad (38)$$

makes m c.c.w encirclements of the origin, where m is the number of unstable poles of $G(j\omega)L(j\omega)$, and

$$G(j\omega) = I_N \frac{p(j\omega)}{1 + p(j\omega)k}.$$

Now consider the case when $G(j\omega)$ and $L(j\omega)$ are stable. Then the stability analysis condition stated above is equivalent to the closed loop system (37) is stable if the Nyquist plot of

$$\prod_{i=1}^{N}[1 + \lambda_i(L(j\omega))G(j\omega)]$$

makes zero encirclements around the origin, which is equivalent to the Nyquist plot of

$$\lambda_i(L(j\omega))G(j\omega)$$

makes zero encirclements around the $-1 + 0j$, for all i.

Furthermore, the above stability condition is equivalent to the closed loop system $H(j\omega)$ is stable if the Nyquist plot of a stable transfer function $G(j\omega)$ does not encircle the points $\frac{-1}{\lambda_i(L(j\omega))}$ for all ω, and $i = 2,\dots n$ (see [12] in the case of static Laplacian matrix). Thus, satisfying this statement means that the closed loop system $H(j\omega)$ is stable and then all the eigenvalues of $H(j\omega)$ are in the left half of the complex plane. Note that the stability of the system does not mean the system will achieve consensus. Thus, we require

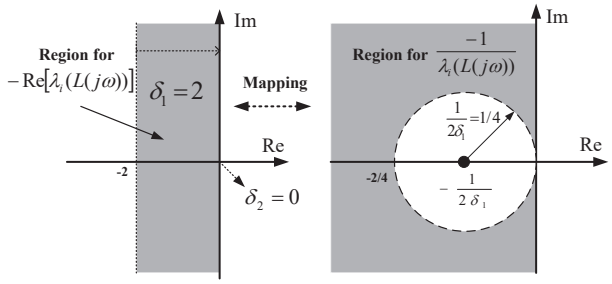


Fig. 5. Left figure: Inclusion region of $-\mathrm{Re}[\lambda_i(L(j\omega))]$. Right figure: Inclusion region of $-\frac{1}{\lambda_i(L(j\omega))}$.

an additional condition for guaranteeing consensus, which is the connectivity of the graph that is associated with the static Laplacian matrix $L(0)$, because under the connectivity condition, the zero eigenvalue of the dynamic Laplacian is distinct and then we can guarantee consensus.

In order to apply the above conditions to designing a controller $k(s)$ such that the closed loop system $H(s)$ is stable, the bounds of the eigenvalues of the dynamic Laplacian must be estimated. These bounds are estimated in Section IV-B and given by (18) and (28).

C. Illustrative Simulations

To illustrate how our approach can be used for the design of controllers, consider the dynamic network Fig. 2 with the dynamic Laplacian $L(s)$ defined by (15) with parameters $Y_{12}(s) = \frac{1}{4s+4}$ and $Y_{13} = Y_{23}(s) = \frac{1}{2s+2}$. Using these parameters, the bounds for the real parts of all non-zero eigenvalues of $L(j\omega)$, $-\mathrm{Re}[\lambda_i(L(j\omega))]$ can be estimated by (18) and (28) as

$$[-\max_{\omega} \mathrm{Re}[\lambda_n(L(j\omega)), -\min_{\omega} \mathrm{Re}[\lambda_2(L(j\omega))]] \subset$$
$$[-2\max_{i} \|\mathrm{Re}[D(j\omega)(i,i)]\|_\infty, \frac{-1}{\|\mathrm{trace}((\mathrm{Re}[L_0(j\omega)])^{-1})\|_\infty}]$$
$$= [-2, 0]. \quad (39)$$

Using the above estimated bounds, the region containing negative of the eigenvalues of the dynamic Laplacian matrix $L(j\omega)$ indicated as the shaded region on the left side of Fig. 5. For the stability analysis described above, we need to define the region of $\frac{-1}{\lambda_i(L(j\omega))}$ for all ω, $i = 1 \ldots N$. Therefore, the bound $-\max_\omega \mathrm{Re}[\lambda_3(L(j\omega)) \geq -4$ is mapped to the exterior of a circle centered at $(-1/4, 0)$ with radius $1/4$. Furthermore, the bound $-\min_\omega \mathrm{Re}[\lambda_2(H(j\omega)) \leq 0]$ is mapped to the left half of the complex plane. Thus, the region of $\frac{-1}{\lambda_i(L(j\omega))}$ is the intersection between the two mapped regions, which is indicated as the shaded region on the right side of Fig. 5. That is a conservative estimate; all the points $-\frac{1}{\lambda_i(L(j\omega))}$ lie in this shaded region. Thus if the Nyquist plot of $G(j\omega)$ does not enter this region, we can be sure consensus occurs.

For this example, assume the nodes of the given dynamic network have a common stable transfer function $p(s) = \frac{\beta}{s+\alpha}$, $\beta = -1, \alpha = 2$. Further, assume a simple static controller $k(s) = k$. Then the stability of the closed loop system

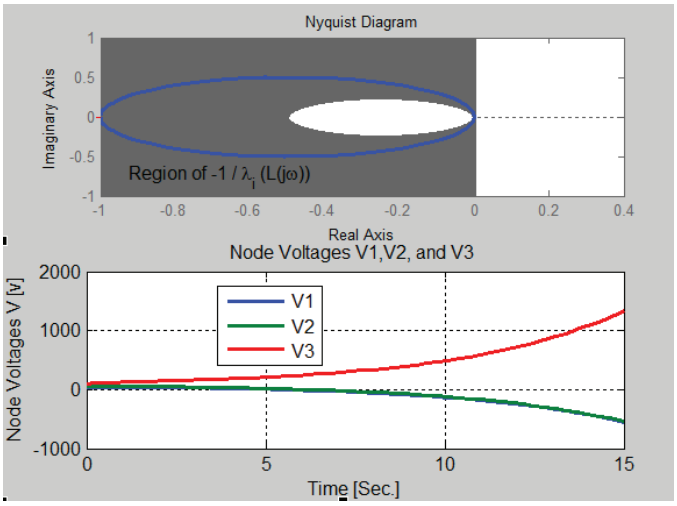


Fig. 6. Upper figure: Nyquist plot of $G(j\omega)$ (blue curve) with the region of $\frac{-1}{\lambda_i(L(j\omega)}$ (shaded region). Lower figure: Simulation of the node voltages in the dynamic network.

discussed above and the estimated bounds can be verified using the Nyquist plot of $G(j\omega)$ with respect to the defined region of $\frac{-1}{\lambda_i(L(j\omega))}$ by considering the following two cases:

Case 1: With a common static controller $k = 1$, the Nyquist plot of $G(j\omega) = I_3 \frac{p(j\omega)}{1+p(j\omega)k}$ (blue curve as shown in the upper side of Fig. 6) does encircle (i.e., intersects) one of the points $\frac{-1}{\lambda_i(L(j\omega))}$. Thus, the closed loop system $H(j\omega)$ in (37) is unstable as shown in the lower side of Fig. 6.

Case 2: If we change the static controller to $k = -0.5$, then the Nyquist plot of $G(j\omega) = I_3 \frac{p(j\omega)}{1+p(j\omega)k}$ (blue curve as shown in the upper side of Fig. 7) does not encircle any point $\frac{-1}{\lambda_i(L(j\omega))}$ because it does not intersect the shaded region. Thus, the closed loop system is stable as shown in the lower side of Fig. 7.

Thus, we can conclude that $k(s)$ stabilizes the closed loop system $H(j\omega)$ if the number of net encirclements of the points $\frac{-1}{\lambda_i(L(j\omega))}$ by the Nyquist plot of the stable transfer function $G(j\omega)$ is zero for all nonzero $\lambda_i(H(j\omega)$. In other words, the stability of the closed-loop system $H(j\omega)$ can be investigated from the locations of the eigenvalues of the dynamic Laplacian matrix $L(j\omega)$, which was estimated in the previous Section, in relation to a Nyquist plot of stable $G(j\omega)$.

In order to design a controller for the proposed dynamic network (identical LTI systems with dynamic edges) that causes all states of nodes to achieve consensus, the controller $k(s)$ must be chosen such that the closed-loop system $H(j\omega)$ is stable or, ideally, reaches consensus. In other words, we must avoid the net encirclement of $-\frac{1}{\lambda_i(L(s))}$ by the Nyquist plot of the stable transfer function $G(j\omega)$ for all nonzero, as was in the second case. For consensus investigation, we require an additional condition, which is the connectedness of $L(0)$ as was previously mentioned, meaning that the zero eigenvalue is distinct.

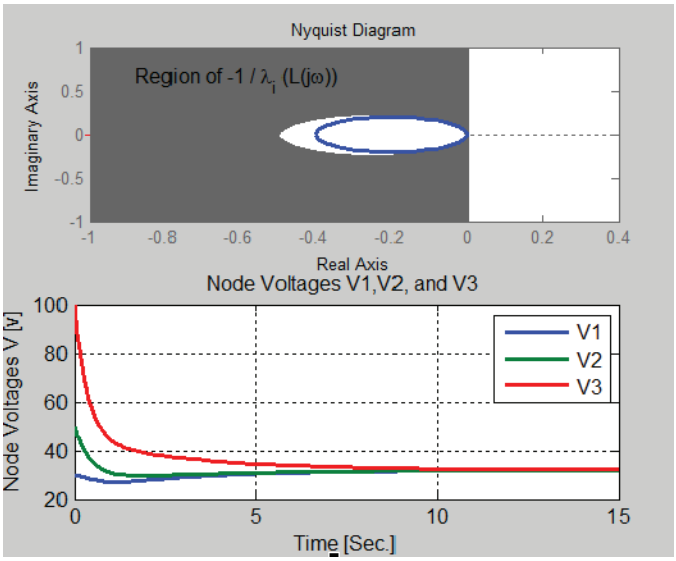


Fig. 7. Upper figure: Nyquist plot of $G(j\omega)$ (blue curve) with the region of $\frac{-1}{\lambda_i(L(j\omega)}$ (shaded region). Lower figure: Simulation of the node voltages in the dynamic network.

VI. Conclusion

In this paper, motivated by several physical examples, including a simple electrical circuit that we presented, we considered dynamic networks whose nodes are integrators and whose edges are SPR transfer functions representing dynamical systems that couple the nodes. We showed that the strictly positive realness of the edges is a sufficient condition for these kinds of dynamic networks to reach a consensus. We developed a method to estimate the bounds of the eigenvalues of the dynamic Laplacian matrix. Our approach used the idea of the *Dynamic Grounded Laplacian*. These bounds were used for a Nyquist stability analysis of a closed loop system comprising an undirected dynamic network subject to distributed control. The region of reciprocal eigenvalues of the Laplacian matrix was defined using the prediction regions of the smallest and largest real parts of the non-zero eigenvalues of the dynamic Laplacian matrix. Using these prediction-regions, the stability conditions of the closed loop system was analyzed with respect to the eigenvalues of the dynamic Laplacian matrix. In future work, we extend these notions to consider the stability and consensus for a dynamic consensus networks with non-identical nodes and dynamic edges.

Acknowledgments

The authors would like to thank Professor Kevin L. Moore and Professor Tyrone L. Vincent at Colorado School of Mines for thier helps and advices.

References

[1] H. Ando, Y. Oasa, I. Suzuki, and M. Yamashita. Distributed memoryless point convergence algorithm for mobile robots with limited visibility. *IEEE Transactions on Robotics and Automation*, 15(5):818–828, 1999.

[2] P. Barooah and J.P. Hespanha. Graph effective resistance and distributed control: Spectral properties and applications. In *Proceedings of 45th IEEE Conference on Decision and Control*, pages 3479 –3485, dec. 2006.

[3] P. Barooah and J.P. Hespanha. Estimation on graphs from relative measurements. *IEEE Control Systems*, 27(4):57 –74, aug. 2007.

[4] J. Cortes, S. Martinez, T. Karatas, and F. Bullo. Coverage control for mobile sensing networks. *IEEE Transactions on Robotics and Automation*, 20(2):243–255, 2004.

[5] M.H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, pages 118–121, 1974.

[6] F. Dorfler and F. Bullo. Synchronization and transient stability in power networks and non-uniform kuramoto oscillators. In *Proceedings of 2010 American Control Conference*, pages 930–937. IEEE, 2010.

[7] J Alexander Fax and Richard M Murray. Information flow and cooperative control of vehicle formations. *IEEE transactions on automatic control*, 49(9):1465–1476, 2004.

[8] J.A. Fax and R.M. Murray. Information flow and cooperative control of vehicle formations. *IEEE Transactions on Automatic Control*, 49(9):1465–1476, 2004.

[9] R.A. Freeman, P. Yang, and K.M. Lynch. Distributed estimation and control of swarm formation statistics. In *Proceedings of 2006 American Control Conference*, pages 7–13. IEEE, 2006.

[10] A. Gattami and R. Murray. A frequency domain condition for stability of interconnected mimo systems. In *Proceedings of the American Control Conference*, volume 4, pages 3723 –3728 vol.4, 30 2004-july 2 2004.

[11] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins Univ Pr, 1996.

[12] S. Hara, T. Hayakawa, and H. Sugata. Stability analysis of linear systems with generalized frequency variables and its applications to formation control. In *Proceedings of 46th IEEE Conference on Decision and Control*, pages 1459 –1466, Dec. 2007.

[13] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[14] F. Lashhab. Dynamic consensus networks: spectral properties, consensus, and control. *Colorado School of Mines. Arthur Lakes Library*, 2012.

[15] R. A. Horn and C. R. Johnson. *Matrix Analysis*. University Press., Cambridge, 1993.

[16] K. Kyo. Decentralized control of multi-agent systems: Theory and applications. *PhD Dissertation, Gwangju Institute of Technology, Gwangju, South Korea*, 2012.

[17] G. Lafferriere, A. Williams, J. Caughman, and JJP Veerman. Decentralized control of vehicle formations. *Systems and Control letters*, 54(9):899–910, 2005.

[18] Fadel Lashhab, Kevin Moore, Tyrone Vincent, Deyuan Meng, and Khalid Kuwairi. Robust H_∞ controller design for dynamic consensus networks. *International Journal of Control*, (just-accepted):1–20, 2017.

[19] J. Lin, AS Morse, and BDO Anderson. The multi-agent rendezvous problem. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, volume 2, pages 1508–1513. IEEE, 2003.

[20] Z. Lin, B. Francis, and M. Maggiore. Necessary and sufficient graphical conditions for formation control of unicycles. *IEEE Transactions on Automatic Control*, 50(1):121–127, 2005.

[21] E. Mallada and Ao Tang. Improving damping of power networks: Power scheduling and impedance adaptation. In *Proceedings of 50th IEEE Conference on Decision and Control and European Control Conference*, pages 7729 –7734, dec. 2011.

[22] K. L. Moore, T. L. Vincent, F. Lashhab, and C. Liu. Dynamic consensus networks with application to the analysis of building thermal processes. *Proceedings of the 18th IFAC World Congress*, 18(1), 2011.

[23] R. Olfati-Saber, J.A. Fax, and R.M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.

[24] R. Olfati-Saber and R.M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004.

[25] R. Olfati-Saber and J.S. Shamma. Consensus filters for sensor networks and distributed sensor fusion. In *Proceedings of 44th IEEE Conference on Decision and Control and European Control Conference*, pages 6698–6703. IEEE, 2005.

[26] A. Papachristodoulou and A. Jadbabaie. Synchronization in oscillator networks: Switching topologies and non-homogeneous delays. In *Proceedings of 44th IEEE Conference on Decision and Control European Control Conference*, pages 5692–5697. IEEE, 2005.

[27] J.W. Polderman and J.C. Willems. *Introduction to Mathematical Systems Theory: a Behavioral Approach*. Springer Verlag, 1998.
[28] O. Egeland R. Lozano, B. Brogliato and B. Maschke. *Dissipative Systems Analysis and Control Theory and Application*. Springer, 2000.
[29] W. Ren, R.W. Beard, and E.M. Atkins. Information consensus in multivehicle cooperative control. *IEEE Control Systems*, 27(2):71–82, 2007.
[30] B.M. Sanandaji, T.L. Vincent, and M.B. Wakin. Exact topology identification of large-scale interconnected dynamical systems from compressive observations. In *Proceedings of American Control Conference*, pages 649–656. IEEE, 2011.
[31] H.G. Tanner, A. Jadbabaie, and G.J. Pappas. Stable flocking of mobile agents part i: Dynamic topology. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, volume 2, pages 2016–2021. IEEE, 2003.
[32] H.G. Tanner, A. Jadbabaie, and G.J. Pappas. Stable flocking of mobile agents, part i: Fixed topology. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, volume 2, pages 2010–2015. IEEE, 2003.
[33] J. Wang and N. Elia. Consensus over networks with dynamic channels. In *Proceedings of 2008 American Control Conference*, pages 2637–2642. IEEE, 2008.
[34] L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. In *Proceedings of Fourth International Symposium on Information Processing in Sensor Networks*, pages 63–70. IEEE, 2005.

APPENDIX

Lemma A.1: Let the dynamic Laplacian matrix $L(j\omega) \in \mathbb{C}^{n\times n}$ and a Dynamic Grounded Laplacian $L_0(j\omega) \in \mathbb{C}^{n_0\times n_0}$ be given, with $n_0 = n-n_g < n$. For any choice of a boundary set of grounded nodes **V0**, for which $L_0(j\omega)$ is a Laplacian that corresponds to a connected graph, a lower bound for the eigenvalues of the Hermitian part of the Dynamic Grounded Laplacian matrix $\lambda_i(H(L_0(j\omega)))$ is given by

$$\lambda_i(H(L_0(j\omega))) \geq \frac{1}{\text{trace}((\text{Re}[L_0(j\omega)])^{-1})}, \tag{40}$$

for all $\omega, i \in 1, 2, ..., n$.

Proof:

It is well known that the sum of the eigenvalues of any matrix is equal to its trace,

$$\sum_{i=1}^{n} \lambda_i(H(L_0(j\omega))^{-1}) = \text{trace}(H(L_0(j\omega))^{-1}). \tag{41}$$

Since $H(L_0(j\omega))^{-1}$ is a real and positive definite matrix (see Lemma 4.1, and Definition 3.1), then

$$\lambda_i(H(L_0(j\omega))^{-1}) > 0, \forall\omega, i = 1, 2, \ldots, n.$$

Thus,

$$\lambda_i(H(L_0(j\omega))^{-1}) \leq \text{trace}(H(L_0(j\omega))^{-1}),$$

$$\lambda_i(H(L_0(j\omega))^{-1}) \leq \text{trace}(\text{Re}[L_0(j\omega)]^{-1}). \tag{42}$$

Since the eigenvalues of $H(L_0(j\omega))$ and $[H(L_0(j\omega))]^{-1}$ are reciprocals of each other, then

$$\lambda_i([H(L_0(j\omega))]^{-1}) = \frac{1}{\lambda_i(H(L_0(j\omega)))}. \tag{43}$$

Combining (42) and (43), we get

$$\lambda_i(H(L_0(j\omega))) \geq \frac{1}{\text{trace}((\text{Re}[L_0(j\omega)])^{-1})}, \tag{44}$$

for all $\omega, i \in 1, 2, ..., n$. ■

Theorem A.2: Interlacing Eigenvalue Theorem [15]: Let A be a Hermitian matrix of order n, and let B be a principal sub matrix of A of order $n-1$. if $\lambda_{min} = \lambda_n \leq \lambda_{n-1} \leq ... \leq \lambda_2 \leq \lambda_1 = \lambda_{max}$ lists the eigenvalues of A and $\mu_n \leq \mu_{n-1} \leq ... \leq \mu_3 \leq \mu_2$ the eigenvalues of B. then

$$\lambda_n \leq \mu_n \leq \lambda_{n-1} \leq \mu_{n-1} \leq ... \leq \lambda_2 \leq \mu_2 \leq \lambda_1. \tag{45}$$

Theorem A.3: Mirsky Theorem [13]: Let $A \in \mathbb{C}^{n\times n}$ be given and set $H(A) = \frac{1}{2}(A + A^H)$. Let $\lambda_i(A)$ and $\lambda_i(H(A))$ denote the eigenvalues of A and $H(A)$, respectively, ordered so that $\text{Re}[\lambda_1(A)] \leq \text{Re}[\lambda_2(A)] \leq ... \leq \text{Re}[\lambda_n(A)]$ and $\lambda_1(H(A)) \leq \lambda_2(H(A)) \leq ... \leq \lambda_n(H(A))$. Then

$$\sum_{i=n-k}^{n} \text{Re}[\lambda_i(A)] \leq \sum_{i=n-k}^{n} \lambda_i(H(A)), \tag{46}$$

$k = 0, 1, ..., n-1$, with equality for $k = n-1$.

Lemma A.4: Let the dynamic Laplacian $L(j\omega) \in \mathbb{C}^{n\times n}$ be given. The Hermitian matrix associated with $L(j\omega)$ is $H(L(j\omega)) = \frac{1}{2}(L(j\omega) + L(j\omega)^H) = \text{Re}[L(j\omega)]$. Let $\lambda_i(L(j\omega))$ and $\lambda_i(H(L(j\omega)))$ denote the eigenvalues of $L(j\omega)$ and $H(L(j\omega))$, respectively, ordered so that $0 = \text{Re}[\lambda_1(L(j\omega))] \leq \text{Re}[\lambda_2(L(j\omega))] \leq ... \leq \text{Re}[\lambda_n(L(j\omega))]$ and $0 = \lambda_1(H(L(j\omega))) \leq \lambda_2(H(L(j\omega)))... \leq \lambda_n(H(L(j\omega)))$. Then the relationship between the real part of the smallest and largest nonzero eigenvalues of $L(j\omega)$ and $H(L(j\omega))$ can be given by

$$\text{Re}[\lambda_2(L(j\omega))] \geq \lambda_2(H(L(j\omega))), \tag{47}$$

and

$$\text{Re}[\lambda_n(L(j\omega))] \leq \lambda_n(H(L(j\omega))). \tag{48}$$

Proof:

Theorem A.3 can be applied for the matrices $L(j\omega)$ and $H(L(j\omega))$ by substituting $k = n-3$ in (46). Thus, we have

$$\sum_{i=3}^{n} \text{Re}[\lambda_i(L(j\omega))] \leq \sum_{i=3}^{n} \lambda_i(H(L(j\omega))), \tag{49}$$

with equality for $k = n-1$

$$\sum_{i=1}^{n} \text{Re}[\lambda_i(L(j\omega))] = \sum_{i=1}^{n} \lambda_i(H(L(j\omega))). \tag{50}$$

For a connected graph, $\text{Re}[\lambda_1(L(j\omega))] = \lambda_1(H(L(j\omega))) = 0$, thus we can rewrite (50) as

$$0 + \text{Re}[\lambda_2(L(j\omega))] + \sum_{i=3}^{n} \text{Re}[\lambda_i(L(j\omega))]$$
$$= 0 + \lambda_2(H(L(j\omega))) + \sum_{i=3}^{n} \lambda_i(H(L(j\omega))).$$

From the above quality and (49), we can conclude

$$\text{Re}[\lambda_2(L(j\omega))] \geq \lambda_2(H(L(j\omega))). \tag{51}$$

Now, substitute $k = 0$ in (46), we get

$$\text{Re}[\lambda_n(L(j\omega))] \leq \lambda_n(H(L(j\omega))). \tag{52}$$

■

Dynamic Consensus Networks: Dynamic Graph Definitions and Controllability Analysis Using the Behavioral Approach

Fadel Lashhab, Member, IEEE
Department of Electrical Engineering and Computer Science
College of Engineering and Architecture
Howard University, Washington DC, USA.
fadel.lashhab@howard.edu

***Abstract*—The focus of this paper is to study a generalization of consensus problems whereby the weights of network edges are no longer modeled as static gains, but instead are represented as dynamic systems coupling the nodes, which might also be more general than an integrator, leading to the notion of dynamic consensus networks. We transform each concept of static graph theory into dynamic terms, out of which a generalized dynamic graph theory naturally emerges. We present a framework for dynamic graphs and dynamic consensus networks. This framework introduces the idea of dynamic degree, adjacency, incident, and Laplacian matrices in a way that naturally extends these concepts from the static case. We consider controllability for dynamic consensus networks. The ideas developed for dynamic graph theory, in conjunction with the behavioral approach, lead to the development of a controllability analysis methodology for dynamic consensus networks. The controllability conditions obtained using the behavioral approach cannot be applied for the general dynamic networks, such as identical LTI nodes with dynamic edges or even in the more general case with heterogeneous nodes. This is because of scalability in such dynamic networks. Thus, we develop controllability conditions based on node and interconnection (edge) parameters that guarantee controllability of the overall dynamic network.**

***Index Terms*—Dynamic graphs, dynamic consensus networks, dynamic Laplacian, Controllability analysis; Behavioral approach.**

I. Introduction

The idea of consensus in networking has received great attention due to its wide array of applications in fields such as robotics, transportation, sensor networking, communication networking, biology, and physics. The focus of this paper is to study a generalization of consensus problems whereby the weights of network edges are no longer static gains, but instead are dynamic systems, leading to the notion of *dynamic consensus networks*.

For the consensus networks, the network topology is static, meaning that there are no dynamics in the interconnections between the nodes ($\lambda_{ij} = constant \geq 0$) and the nodes are assumed to be integrators [1]. Thus, static consensus problems can be written in the time domain for each node $i = 1, 2, ..., n$ as

$$\dot{x}_i = \sum_{j \in \mathcal{N}_i} \lambda_{ij}(x_j(t) - x_i(t)). \quad (1)$$

The continuous time linear consensus protocol (1) can be written in matrix form as:

$$\dot{x}(t) = -Lx(t), \quad (2)$$

where $x(t) = [x_1(t), x_2(t), ..., x_n(t)]^T$ and L, the graph's Laplacian matrix $L = [l_{ij}]$, is defined by

$$l_{ij} = \begin{cases} \sum_{j \in \mathcal{N}_i} \lambda_{ij} & i = j \\ -\lambda_{ij} & i \neq j \text{ and } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

For the multi-agent consensus problem, suppose that N agents evolve their individual beliefs $x_i \in \mathbb{R}^1$ about a so-called global consensus variable x using communications with their nearest neighbors according to the consensus protocol (1). A key result is that the solution of $\dot{x}(t) = -Lx(t)$ gives $x_i \rightarrow x^*$ if the static graph is connected [1]. This specific fact has been the basis of much of the literature related to consensus problems.

The consensus protocol such as in (1) and its variants have been studied extensively in the literature and have been applied in a number of areas, notably for formation motion control when the agents are mobile. However, the consensus paradigm is restrictive in a number of ways. Notice that we have interpreted the consensus problem as having integrating nodes and static weights. One might ask: what if the weights were also transfer functions? What if the nodes are more than integrators? In next section we have considered this question, motivated by modeling heat transfer in buildings [2]. In this work we have considered *dynamic consensus networks*, by modeling a thermal process in a building as a directed dynamic graph. We consider the hypothetical four-room building in which each room has several neighbors with which it is interconnected. Pathways include walls, doors,

978-1-7281-9615-2/20$31.00

and windows. The interconnection between the two rooms is a wall, which is represented analogously by an electrical circuit with three resistors and two capacitors known as 3R2C model in the literature. The corresponding graph for this example (we consider dynamic network with integrator nodes $P_i(s) = \frac{1}{s}, \forall i = 1...N$ and dynamic edges) is shown in Fig. 1. In developing a model for this system (see [2]), the sum of the energy losses through all pathways connected to a node, resulting in:

$$sT_i(s) = Q_i^{in}(s) - \sum_{j\in\mathcal{N}_i} \left[\lambda_{ij}^S(s)T_i(s) - \lambda_{ij}^C(s)T_j(s)\right]. \quad (4)$$

Thus, if we define the vectors

$$T(s) = \begin{bmatrix} T_1(s) & T_2(s) & \dots & T_N(s) \end{bmatrix}^T,$$
$$Q^{in}(s) = \begin{bmatrix} Q_1^{in}(s) & Q_2^{in}(s) & \dots & Q_N^{in}(s) \end{bmatrix}^T,$$

then we can write

$$sT(s) = Q^{in}(s) - L(s)T(s), \quad (5)$$

where the matrix $L(s) = [L_{ij}(s)]$ is given as:

$$L_{ij}(s) = \begin{cases} \sum_{j\in\mathcal{N}_i} \lambda_{ij}^S(s) & i = j \\ -\lambda_{ij}^C(s) & i \neq j \text{ and } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

where $\lambda_{ij}^S(s)$, $-\lambda_{ij}^C(s)$, and more details for modeling of the above results are defined in [2]. We will refer to $L(s)$ defined in this way as a *dynamic Laplacian matrix*. For the graph topology shown in Fig. 1, the dynamic Laplacian matrix has the form shown in (15). We have shown that when the weight matrices $\lambda_{ij}(s)$ satisfy certain assumptions, $-L(s)$ can be viewed as a *dynamic interconnection matrix*, allowing the demonstration of consensus. The graph shown in Fig. 1 will be referred-to as a *dynamic graph* or a *dynamic consensus network*.

For the autonomous case ($Q_i^{in}(t) = 0$) and assume for the generality ($x_i(t) = T_i(t), \lambda_{ij}^S(s) = \lambda_{ij}^C(s) = \lambda_{ij}(s)$), the dynamic of each node in (4) can then be written in time domain as:

$$\dot{x}_i(t) = \sum_{j\in\mathcal{N}_i} \lambda_{ij}(t) * [x_j(t) - x_i(t)]. \quad (7)$$

By analogy with the static case (1), (7) is referred to as the *dynamic consensus protocol*. Motivated by the single integrator dynamic consensus network (7), we consider more complicated dynamics ($p_i(s)$ for the node i, $i = 1, 2, ..., N$) than simple integrator nodes as shown in Fig. 1. For this dynamic network, we suppose that the node dynamic is given by

$$x_i(s) = p_i(s) \sum_{j\in\mathcal{N}_i} \lambda_{ij}(s)[x_j(s) - x_i(s)]. \quad (8)$$

We call a dynamic network, such as that defined in (8), a dynamic consensus network with heterogeneous nodes and dynamic edges.

Recently, we consider another example motivated a generalization of the static consensus problem (1), modeling of the load frequency control (LFC) network of an electrical power grid as dynamic consensus network [3]. We consider the following network:

$$Y_i(s) = \frac{1}{s} \sum_{j\in\mathcal{N}_i} G_i(s)a_{ij}(Y_j(s) - Y_i(s)), \quad (9)$$

$i = 1, \dots, N$, which can be viewed as a single-integrator consensus network with dynamic interconnection coefficients $G_i(s)a_{ij}$. In the LFC network of the grid, the output of each individual system is the phase of its voltage, which is the integration of the angular velocity. The interconnection is power exchanges among the individual systems through transmission lines, which are dependent on phase differences.

The paper is organized as follows: Section II presents a framework for dynamic graphs and dynamic consensus networks. This framework introduces the idea of dynamic degree, adjacency, incident, and Laplacian matrices in a way that naturally extends these concepts from the static case. From this, we can easily define equivalent concepts of dynamic interconnection matrices and dynamic consensus networks. In Section III, we use the established aspects and properties of the defined dynamic graph theory in conjunction with the behavioral approach to develop a controllability-analysis methodology for dynamic networks. Because of scalability in such dynamic networks, the controllability conditions based on node and interconnection (edge) parameters that guarantee controllability of the overall dynamic network were developed in Section 4.

II. Dynamic Graphs Definitions

The previous work [2], [3] showed how two different phenomena can be modeled in the form of a graph whose edges are transfer functions (i.e., dynamic systems). In this section, from the motivation of the models developed in the previous work, we generalize all the typical notations from (static) graph theory to the dynamic case. The development here parallels the notations in static graph.

Consider the example of a directed, dynamic graph shown in Fig. 1. Such graphs can be described as a set of nodes (or vertices) $\mathcal{N} = \{n_i\}$ connected by a set of edges $\mathcal{E}(s) = \{(n_i, n_j) : n_i, n_j \in \mathcal{N}\}$. Each edge is modeled as a transfer function $\lambda_{ij}(s)$. More formally, we can say that each edge $\lambda_{ij}(s) \in \mathcal{R}(s)$, where $\mathcal{R}(s)$ denotes the set of all complex-valued functions analytic in the open right-half complex plane (real rational functions). We assume there are no self-loops associated with any node. If there is an edge between nodes n_i and n_j, we say these nodes are adjacent (or neighbors). We denote the neighbors of node n_i as $\mathcal{N}_i = \{j : (n_i, n_j) \in \mathcal{E}(s)\}$. A path between two nodes is a sequence of edges by which it is possible to move along the sequence of arcs from one of the nodes to the other. If there is at least one node that

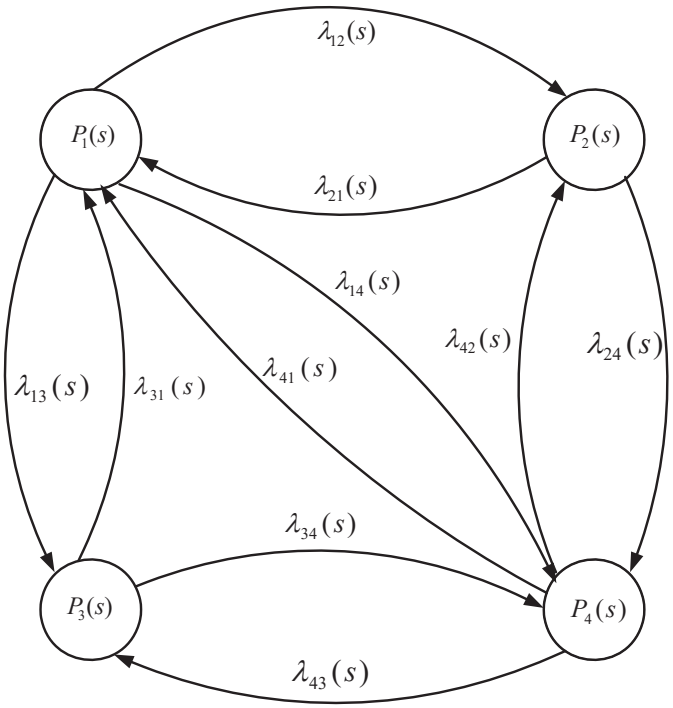


Fig. 1. Directed-Dynamic Graph.

has at least one path to every other node, the graph is said to be connected.

Later, we also view a node as implementing a transfer function that produces the node variable ($P_i(s)$ for $i = 1, 2, ..., N$, where N is number of nodes in the dynamic graph) by processing the incoming and outgoing flows.

As in the static case, the edges $e_{ij}(s)$ are ordered by the edge originating from node $P_i(s)$, known as the tail node, and terminating at node $P_j(s)$, known as the head node, which can be identified by the presence of an arrowhead. If the dynamics of the nodes are different, we call such dynamic networks heterogeneous. If the nodes have the same dynamics, the network is called homogeneous. A special case is when we have integrator nodes with dynamic edges.

Each node $P_i(s)$ in a directed dynamic graph, such as Fig. 1, is associated with a dynamic degree $\nu_i(s)$ representing the total sum of the dynamic edge weights that are connected to the node i. More specifically, each node has a dynamic in-degree $\nu_i^{in}(s)$ and a dynamic out-degree $\nu_i^{out}(s)$ representing the sum of the dynamic edge weights of the incoming and outgoing edges, respectively. Clearly $\nu_i(s) = \nu_i^{in}(s) + \nu_i^{out}(s)$. From these dynamic degree definitions, we can define three different dynamic degree matrices:

1) The dynamic in-degree matrix $D^{in}(s) = diag(\nu_i^{in}(s))$.
2) The dynamic out-degree matrix $D^{out}(s) = diag(\nu_i^{out}(s))$.
3) The dynamic degree matrix $D(s) = diag(\nu_i(s))$.

Notice that $D(s) = D^{in}(s) + D^{out}(s)$.

If a dynamic edge $e_{ij}(s)$ exists between two nodes i and j, these nodes are considered to be adjacent and are known as neighbors, and are denoted for a node $P_i(s)$ by $\mathcal{N}_i = \{j : (n_i, n_j) \in \mathcal{E}(s)\}$. As before, neighbors can be distinguished based upon whether they are associated with incoming or outgoing arcs. Thus, we can define three dynamic adjacency matrices:

1) The incoming dynamic adjacency matrix $A^{in}(s) = [a_{ij}^{in}(s)]$, is defined by

$$a_{ij}^{in}(s) = \begin{cases} \sum_{j\in\mathcal{N}_i} e_{ij}(s); (\ coming\ into\ n_i \text{ from } n_j) & \text{if } i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

2) The dynamic outgoing adjacency matrix $A^{out}(s) = [a_{ij}^{out}(s)]$, is defined by

$$a_{ij}^{out}(s) = \begin{cases} \sum_{j\in\mathcal{N}_i} e_{ij}(s); (\ going\ out\ from\ n_i \text{ into } n_j) & \text{if } i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

3) The dynamic adjacency matrix $A(s) = [a_{ij}(s)]$, is defined by

$$a_{ij}(s) = \begin{cases} \text{total} \sum_{\text{j}\in\mathcal{N}_\text{i}} \text{e}_\text{ij}(\text{s}); (\text{ between } n_i \text{ and } n_j) & \text{if } i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $A(s) = A^{in}(s) + A^{out}(s)$.

Another type of dynamic matrix is the dynamic incident matrix. For the incoming incident matrix, we will define two incident matrices: one indicates the direction of the edges connected to a node, where for node $P_i(s)$ the edge $e_{ij}(s)$ is given a value based upon being disconnected, incoming, or outgoing and denoted as B_S^{in} where "S" refers to static. Another matrix captures the transfer functions of the edges and is denoted as $B_D^{in}(s)$ where "D" refers to dynamic. Thus, the dynamic and static incoming incident matrices are defined as $B_D^{in}(s) = [b_{ij-D}^{in}(s)]$, $B_S^{in} = [b_{ij-S}^{in}]$, where,

$$b_{ij-D}^{in}(s) = \begin{cases} +\lambda_{ij}(s) & \text{if arc } j \text{ enters node } n_i \\ 0 & \text{otherwise.} \end{cases}$$

$$b_{ij-S}^{in} = \begin{cases} +1 & \text{if arc } j \text{ enters node } n_i \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, we can define a dynamic and static outgoing incident matrices for a dynamic graph by, $B_D^{out}(s) = [b_{ij-D}^{out}(s)]$, $B_S^{out} = [b_{ij-S}^{out}]$, where,

$$b_{ij-D}^{out}(s) = \begin{cases} -\lambda_{ij}(s) & \text{if arc } j \text{ leaves node } n_i \\ 0 & \text{otherwise.} \end{cases}$$

$$b_{ij-S}^{out} = \begin{cases} -1 & \text{if arc } j \text{ leaves node } n_i \\ 0 & \text{otherwise.} \end{cases}$$

Also, we can define the dynamic and static incident matrices for a directed dynamic graph as $B_D(s) = [b_{ij-D}(s)]$, $B_S = [b_{ij-S}]$ where

$$b_{ij-D}(s) = \begin{cases} -\lambda_{ij}(s) & \text{if arc } j \text{ leaves node } n_i \\ +\lambda_{ij}(s) & \text{if arc } j \text{ enters node } n_i \\ 0 & \text{otherwise.} \end{cases}$$

$$b_{ij-S} = \begin{cases} -1 & \text{if arc } j \text{ leaves node } n_i \\ +1 & \text{if arc } j \text{ enters node } n_i \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $B_D(s) = B_D^{in}(s) + B_D^{out}(s)$ and $B_S = B_S^{in} + B_S^{out}$.

Having derived dynamic degree, adjacency, and incident matrices, we can now give the dynamic equivalent of the static Laplacian matrix. The dynamic Laplacian matrix has spectral properties that indicate many facts about a graph. An undirected dynamic graph, one whose dynamic edges are not directionally-fixed, has a corresponding dynamic Laplacian matrix defined with no ambiguities by $L(s) = D(s) - A(s)$. More specifically, the dynamic Laplacian matrix is defined as $L(s) = [l_{ij}(s)]$, where

$$l_{ij}(s) = \begin{cases} \sum_{j \in \mathcal{N}_i} \lambda_{ij}(s) & i = j \\ -\lambda_{ij}(s) & i \neq j \text{ and } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}. \quad (10)$$

We can also define the dynamic Laplacian matrix of an undirected dynamic graph as

$$L(s) = BD(s)B^T = D(s) - A(s), \quad (11)$$

where $D(s) \in \mathbb{C}^{m \times m}$ is the dynamic degree matrix formed by the dynamic degree of the m edges, $B \in \mathbb{R}^{n \times m}$ is the static incident matrix that captures the orientations of the edges, and $A(s) \in \mathbb{C}^{n \times n}$ is the dynamic adjacency matrix.

In the static case, the definition of the Laplacian matrix for a directed graph requires adopting either an incoming or outgoing convention. Likewise, we define the dynamic Laplacian matrix utilizing the dynamic degree and dynamic adjacency matrices distinguishing between incoming and outgoing conventions. Examples include: $L^{in}(s) = D^{in}(s) - A^{in}(s)$ and $L^{out}(s) = D^{out}(s) - A^{out}(s)$. With these definitions, $L(s) = L^{in}(s) + L^{out}(s)$. Particular caution must be taken in noting that while $L = BB^T$ in the static case, $L(s) \neq B(s)B(s)^T$, $L^{in}(s) \neq B^{in}(s)B^{in}(s)^T$ and $L^{out}(s) \neq B^{out}(s)B^{out}(s)^T$. To overcome this, in the sequel, we will use $L^{out}(s)$, where the outgoing dynamic Laplacian matrix can be defined using the incident matrices as follow:

$$L^{out}(s) = D^{out}(s) - A^{out}(s), \quad (12)$$

where $D^{out}(s)$ and $A^{out}(s)$ are the dynamic-outgoing degree and adjacency matrices, respectively. These matrices can be defined in a static case using the incident matrices as $D^{out} = B^{out}B^{out^T}$ and $A^{out} = -B^{out}B^{in^T}$. For the dynamic graphs, $D^{out}(s) \neq B^{out}(s)B^{out}(s)^T$ because the product of the dynamic incident matrices results in a matrix where its elements are the square of the edge dynamics (i.e., $\lambda_{ij}(s)^2$). To overcome this problem, we use the dynamic and static incident matrices (one will capture the edges orientation and the other will capture the edges dynamic) for defining the dynamic degree and adjacency matrices. Thus, the outgoing dynamic degree and adjacency matrices can be defined as:

$$D^{out}(s) = B_D^{out}(s)B_S^{out^T},$$
$$A^{out}(s) = -B_D^{out}(s)B_S^{in^T}. \quad (13)$$

By combining (12) and (13), the outgoing, dynamic Laplacian matrix $L^{out}(s)$ can be defined as

$$\begin{aligned} L^{out}(s) &= B_D^{out}(s)B_S^{out^T} + B_D^{out}(s)B_S^{in^T} \\ &= B_D^{out}(s)(B_S^{out^T} + B_S^{in^T}) = B_D^{out}(s)B_S^T. \end{aligned} \quad (14)$$

A similar definition can be given for $L^{in}(s)$.

To illustrate, for the example shown in Fig. 1, the associated dynamic degree, adjacency, and Laplacian matrices are given by (15). Note that all $\lambda_{ij}(s)$ in (15) are transfer functions that describe the interconnections (edges) between the nodes.

III. Controllability Analysis for Dynamic Consensus Networks Using the Behavioral Approach

In this section, we use the established aspects and properties of the defined dynamic graph theory in conjunction with the behavioral approach to develop a controllability-analysis methodology for dynamic networks. Additionally, we demonstrate the application of the dynamic interconnection topology (dynamic Laplacian matrix) in analyzing the interconnections of nodes for the purpose of establishing controllability.

Controllability analysis for consensus networks with static topology has already been studied by several researchers. Most of these studies have investigated the effect of the static topology on the controllability of consensus networks. The authors of [4] introduced a graph-theoretic characterization of the structural controllability of static networks with a single leader and showed that a static network with switching topology is structurally-controllable if and only if the union graph of the underlying static topologies is connected. In [5], controllability was investigated using the size of the graph and its connectivity. Controllability for leader-based, multi-agent systems was analyzed in [6] based on connectivity and the null space of the incidence matrices of the leader and followers. Controllability using the graph symmetry, automorphism, and equitable-portions properties was addressed in [7]. In this section, we analyze the dynamic interconnection topologies of the proposed dynamic networks for the purpose of establishing controllability using the behavioral approach to control [8]. We first describe the behavioral approach and then discus the controllability of the two room example using this approach. We then give the general result.

$$D^{out}(s) = \begin{bmatrix} \lambda_{12}(s)+\lambda_{13}(s)+\lambda_{14}(s) & 0 & 0 & 0 \\ 0 & \lambda_{21}(s)+\lambda_{24}(s) & 0 & 0 \\ 0 & 0 & \lambda_{31}(s)+\lambda_{34}(s) & 0 \\ 0 & 0 & 0 & \lambda_{41}(s)+\lambda_{42}(s)+\lambda_{43}(s) \end{bmatrix},$$

$$A^{out}(s) = \begin{bmatrix} 0 & \lambda_{12}(s) & \lambda_{13}(s) & \lambda_{14}(s) \\ \lambda_{21}(s) & 0 & 0 & \lambda_{24}(s) \\ \lambda_{31}(s) & 0 & 0 & \lambda_{34}(s) \\ \lambda_{41}(s) & \lambda_{42}(s) & \lambda_{43}(s) & 0 \end{bmatrix},$$

$$L^{out}(s) = \begin{bmatrix} \lambda_{12}(s)+\lambda_{13}+\lambda_{14}(s) & -\lambda_{12}(s) & -\lambda_{13}(s) & -\lambda_{14}(s) \\ -\lambda_{21}(s) & \lambda_{21}(s)+\lambda_{24}(s) & 0 & -\lambda_{24}(s) \\ -\lambda_{31}(s) & 0 & \lambda_{31}(s)+\lambda_{34}(s) & -\lambda_{34}(s) \\ -\lambda_{41}(s) & -\lambda_{42}(s) & -\lambda_{43}(s) & \lambda_{41}(s)+\lambda_{42}(s)+\lambda_{43}(s) \end{bmatrix}. \quad (15)$$

A. Behavioral Approach

The behavior of a dynamic system is the collection of all possible time trajectories in the system [9], [10]. In other words, the behavior is a family of trajectories, rather than a transfer function. The behavior of a dynamic system is defined as the set of solutions of a system of linear, constant-coefficient differential equations. For example, consider a system described by the following set of differential equations:

$$\begin{aligned} &w_1 + 2\frac{d}{dt}w_1 + \frac{d^2}{dt^2}w_1 - w_2 - \frac{d}{dt}w_2 = 0; \\ &-w_1 - \frac{d}{dt}w_1 + 7w_2 + 5\frac{d}{dt}w_2 + \frac{d^2}{dt^2}w_2 - 6w_3 - \\ &4\frac{d}{dt}w_3 = 0; \end{aligned} \quad (16)$$

The dynamic system (16) can be written in matrix form as:

$$M\left(\frac{d}{dt}\right)w = 0, \quad (17)$$

where

$$M(\frac{d}{dt}) = \begin{bmatrix} \frac{d^2}{dt^2}(\cdot)+2\frac{d}{dt}(\cdot)+1 & -(\frac{d}{dt}(\cdot)+1) & 0 \\ -(\frac{d}{dt}(\cdot)+1) & \frac{d^2}{dt^2}(\cdot)+5\frac{d}{dt}(\cdot)+7 & -4\frac{d}{dt}(\cdot)-6 \end{bmatrix},$$

and $w = [w_1, w_2, w_3]^T$. We can also describe (16) using polynomial notation as

$$\begin{bmatrix} s^2+2s+1 & -(s+1) & 0 \\ -(s+1) & s^2+5s+7 & -4s-6 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = 0, \quad (18)$$

From the above equations, we can define the behavior of the system (16) as all possible sets of solutions (w_1, w_2, and w_3) that satisfy (17) or (18). We call equation (17) or (18) a kernel representation of the dynamic system (16) [8].

We can use the behavioral system theory notation $M(\frac{d}{dt})$ to describe the relationships between the variables at a node can be represented by a kernel operator made up of differential equations. For more details see [11]. To illustrate, given a node with the transfer function description:

$$T_i(s) = -\frac{10}{s+1}q_{i1}(s) - \frac{s+4}{s+2}q_{i2}(s) + \frac{s+5}{s+1}Q_i^{in}(s),$$

which could be rewritten as

$$T_i(s) + \frac{10}{s+1}q_{i1}(s) + \frac{s+4}{s+2}q_{i2}(s) - \frac{s+5}{s+1}Q_i^{in}(s) = 0,$$

we can describe this as $M(\frac{d}{dt})x(t) = 0$, where $x(t) = [T_i(t)\ q_{i1}(t)\ q_{i2}(t)\ Q_i^{in}(t)]^T$ and $M(\frac{d}{dt})$ is defined by (19).

With this notation, an edge can be represented as $M^{i\leftrightarrow j}(\frac{d}{dt})$, as is done in considering the two node example shown in Fig. 2. The transfer function in the connection element is equivalently:

$$\begin{bmatrix} -(\frac{d}{dt}(\cdot)+3) & 0 & \frac{d}{dt}(\cdot)+1 & 10 \\ 0 & -(\frac{d}{dt}(\cdot)+3) & 10 & \frac{d}{dt}(\cdot)+2 \end{bmatrix}\begin{bmatrix} q_1 \\ q_2 \\ T_1 \\ T_2 \end{bmatrix} = 0. \quad (20)$$

The kernel representation for the system can then be defined by:

$$M\left(\frac{d}{dt}\right)x(t) = 0. \quad (21)$$

B. Controllability using the Behavioral Representation

Definition 3.1: [8] The system (21) is *controllable* if for any two trajectories $x_1(t)$ and $x_2(t)$ (as shown in Fig. 3) satisfying (21) there exists a time $T \geq 0$ and a third trajectory $x_3(t)$ satisfying (21) such that

$$x_3(t) = \begin{cases} x_1(t) & t \leq 0 \\ x_2(t-T) & t \geq T \end{cases}$$

In a controllable system, it is possible to utilize freely-assignable variables to switch between any legal past and future trajectories with some delay T as shown in Fig. 3 . Thus, for a controllable system it is possible to reach any legal future trajectory regardless of its present state or using the transition trajectory. In [12], controllability was described using the kernel representation:

Theorem 3.1:

The system (21) is controllable if and only if the rank of $M(s)$ is the same for all $s \in \mathbb{C}$.

$$M(\frac{d}{dt}) = [(\frac{d^2}{dt^2}(\cdot) + 3\frac{d}{dt}(\cdot) + 2) \quad 10(\frac{d}{dt}(\cdot) + 2) \quad (\frac{d^2}{dt^2}(\cdot) + 5\frac{d}{dt}(\cdot) + 4) \quad -(\frac{d^2}{dt^2}(\cdot) + 7\frac{d}{dt}(\cdot) + 10)] \tag{19}$$

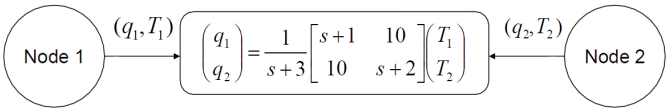


Fig. 2. Connection transfer function example.

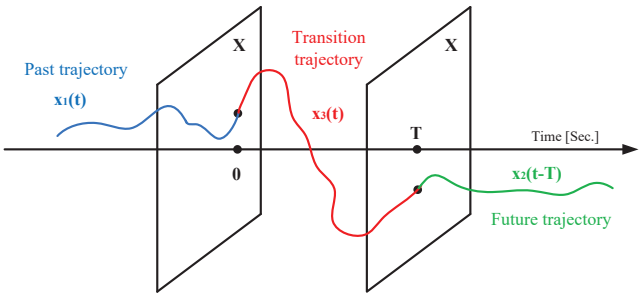


Fig. 3. Behavioral controllability.

C. Controllability of Two Rooms: An Example

As an example, we investigate controllability of the two room example. Our problem is first formulated by defining the behavioral representation of the example containing two rooms (For more details in the modeling of two rooms example see our previous work in [2]). Fig. 4 depicts what is called a 3R2C model. We identify a room i as a node with node variable T_i, the lumped room temperature, Q_i^{in}, the input heat flow (a manipulated variable, not shown), and q_{ij}, the heat flow out of the room through walls or doors or windows (of course with this convention, if $q_{ij} < 0$ then this is heat flow into the room). The parameter C_i^r is the thermal capacity (mass) of the room i.

The interconnection between the two rooms is a wall, which is represented analogously by an electrical circuit with three resistors and two capacitors, a simplification of the model in Fig. 4. As shown in [2], the heat flows in Fig. 4 can be written

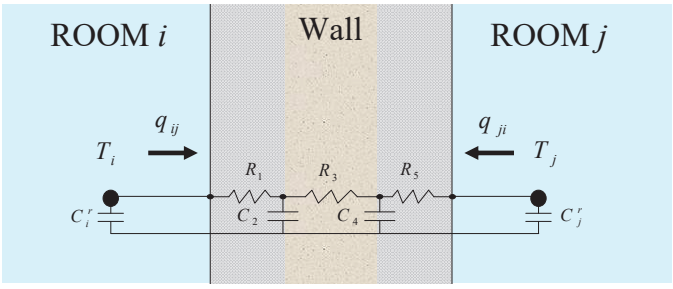


Fig. 4. Two rooms connected by a wall using the 3R2C model.

as:

$$\begin{bmatrix} q_{ij} \\ q_{ji} \end{bmatrix} = \frac{1}{B_{ij}(s)}\begin{bmatrix} A_{ij}(s) & -D_{ij}(s) \\ -D_{ij}(s) & A_{ji}(s) \end{bmatrix}\begin{bmatrix} T_i \\ T_j \end{bmatrix}, \tag{22}$$

where $A_{ij}(s), A_{ji}(s), B_{ij}(s)$, and $D_{ij}(s)$ are defined in [2].

From Fig. 4, the nodal equation can be written as:

$$C_i^r \frac{dT_i}{dt} = Q_i^{in} - q_{ij} \tag{23}$$

Combining (22) and (23) gives:

$$\begin{bmatrix} C_i^r s & 0 \\ 0 & C_j^r s \end{bmatrix}\begin{bmatrix} T_i \\ T_j \end{bmatrix} = \begin{bmatrix} Q_1^{in} \\ Q_2^{in} \end{bmatrix} - \frac{1}{B_{ij}(s)}\begin{bmatrix} A_{ij}(s) & -D_{ij}(s) \\ -D_{ij}(s) & A_{ji}(s) \end{bmatrix}\begin{bmatrix} T_i \\ T_j \end{bmatrix} \tag{24}$$

From (24), the kernel representation of the two rooms example can be expressed by (25).

Using Theorem 3.1 and the structures of the system's matrices, the following theorem defines the controllability of the two rooms example:

Theorem 3.2: Suppose for all i, j, the roots of $B_{ij}(s)$ (in 24) are isolated. Then system (25) is uncontrollable if and only if there exists a root λ of $B_{ij}(s)$ such that

$$\begin{bmatrix} A_{ij}(\lambda) & -D_{ij}(\lambda) \\ -D_{ij}(\lambda) & A_{ji}(\lambda) \end{bmatrix} \tag{26}$$

is not full row rank.

Proof:

Based on Theorem 3.1, the system (25) is controllable if and only if the rank of $M(s)$ is the same for all $s \in \mathbb{C}$. Suppose λ is the root of $B_{ij}(s)$ ($B_{ij}(\lambda) = 0$), then we can prove the theorem by considering the following two cases:

Case 1: λ not a root of $B_{ij}(s)$: The system is controllable because $B_{ij}(\lambda) \neq 0$, thus the rank of $M(s)$ is two.

Case 2: λ a root of $B_{ij}(s)$: In this case $B_{ij}(\lambda) = 0$, so $M(s)$ becomes

$$M(\lambda) = \begin{bmatrix} A_{ij}(\lambda) & -D_{ij}(\lambda) & 0 & 0 \\ -D_{ij}(\lambda) & A_{ji}(\lambda) & 0 & 0 \end{bmatrix}. \tag{27}$$

Now we can see from the last equation that the system is uncontrollable if and only if the non-zeros block matrix

$$\begin{bmatrix} A_{ij}(\lambda) & -D_{ij}(\lambda) \\ -D_{ij}(\lambda) & A_{ji}(\lambda) \end{bmatrix} \tag{28}$$

is not full row rank. ■

A similar result is also possible with repeated roots where the rank condition would have to be verified for terms that share roots of the same or higher multiplicity.

$$\underbrace{\begin{bmatrix} sC_i^r B_{ij}(s) + A_{ij}(s) & -D_{ij}(s) & -B_{ij}(s) & 0 \\ -D_{ij}(s) & sC_j^r B_{ij}(s) + A_{ji}(s) & 0 & -B_{ij}(s) \end{bmatrix}}_{M(s)|_{s=\frac{d}{dt}}} \begin{bmatrix} T_i \\ T_j \\ Q_i^{in} \\ Q_j^{in} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$M\left(\frac{d}{dt}\right) x(t) = 0. \tag{25}$$

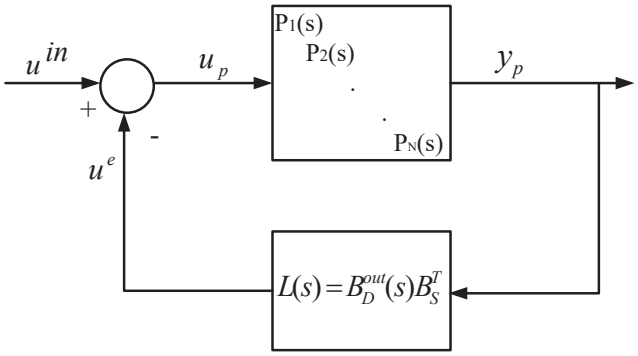


Fig. 5. Block diagram of the dynamic network with heterogeneous nodes and dynamic edges.

IV. General Controllability of Dynamic Consensus Networks

The controllability conditions obtained in the previous section cannot be applied for the general dynamic networks, such as identical LTI nodes with dynamic edges or even in the more general case with heterogeneous nodes. This is because of scalability in such dynamic networks. For example, we have a similar result for the four rooms example but this result will be difficult to apply when we consider more rooms or nodes in a dynamic network. Thus, we seek to develop controllability conditions based on node and interconnection (edge) parameters that guarantee controllability of the overall dynamic network.

The block diagram for the general dynamic networks (heterogeneous nodes with dynamic edges as proposed in (8)) can be depicted as shown in Fig. 5. To propose our task, our problem is first formulated by considering a dynamic consensus network consisting of N heterogeneous LTI nodes $P_i(s)$ for all $i = 1, \ldots, N$ modeled by the following state-space equations:

$$\begin{aligned} \dot{x}_{p,i}(t) &= A_{p,i}x_{p,i}(t) + B_{p,i}u_{p,i}(t) \\ y_{p,i}(t) &= C_{p,i}x_{p,i}(t), \end{aligned} \tag{29}$$

where $x_{p,i}(t) \in \mathbb{R}^{n_i}$, $u_{p,i}(t) \in \mathbb{R}^{m_i}$, $y_{p,i}(t) \in \mathbb{R}^{p_i}$ denote the state, input, and output, respectively, of node i for $i = 1, \ldots, N$ and $A_{p,i}, B_{p,i}, C_{p,i}$ are constant matrices with appropriate dimensions. If we define the vectors $x_p(t) = [x_{p,1}(t)^T, \ldots, x_{p,N}(t)^T]^T$, $u_p(t) = [u_{p,1}(t)^T, \ldots, u_{p,N}(t)^T]^T$, and $y_p(t) = [y_{p,1}(t)^T, \ldots, y_{p,N}(t)^T]^T$ we can write (29) in a matrix form as:

$$\begin{aligned} \dot{x}_p(t) &= A_p x_p(t) + B_p u_p(t) \\ y_p(t) &= C_p x_p(t), \end{aligned} \tag{30}$$

where, $x_p(t) \in \mathbb{R}^n$, $u_p(t) \in \mathbb{R}^m$, $y_p(t) \in \mathbb{R}^p$, $n = \sum_{i=1}^N n_i$, $m = \sum_{i=1}^N m_i$, $p = \sum_{i=1}^N p_i$, $A_p = \text{diag}(A_{p,i})$, $B_p = \text{diag}(B_{p,i})$, and $C_p = \text{diag}(C_{p,i})$.

Note that the input vector into the nodes $u_p(t)$ (see Fig. 5) is given by:

$$u_p(t) = u(t)^{in} + u^e(t), \tag{31}$$

where $u^{in}(t) = [u_1^{in}(t)^T, ..., u_N^{in}(t)^T]^T$ is the input vector from the environment to the nodes in the dynamic network and $u^e(t) = [u_1^e(t)^T, \ldots, u_N^e(t)^T]^T$ is the input vector from the dynamic topology (edges) to the nodes. Assume that the interconnections (edges) of the dynamic topology are modeled by transfer functions $e_{ij}(s) = \lambda_{ij}(s)$. Thus, the dynamic consensus protocol that describes the dynamic topology is given in the frequency domain by the following equations (see (8):

$$u_i^e(s) = \sum_{j \in \mathcal{N}_{p,i}} \left[\lambda_{ij}(s)(y_{p,j}(s) - y_{p,i}(s))\right], \forall i = 1 \ldots N. \tag{32}$$

Note that the outgoing convention has been used for describing the dynamic topology. The dynamic consensus protocol (32) can be expressed in a matrix form as:

$$u^e(s) = -L^{out}(s) y_p(s), \tag{33}$$

where $L^{out}(s)$ is the outgoing dynamic Laplacian matrix defined in (14) as

$$L^{out}(s) = B_D^{out}(s) B_S^{out^T} + B_D^{out}(s) B_S^{in^T} = B_D^{out}(s) B_S^T, \tag{34}$$

where $B_D^{out}(s)$ and B_S are the dynamic outgoing incident matrix and dynamic incident matrix, respectively, and these matrices are defined in Section II for the example shown in Fig. 1.

Combining (33) and (34), we get

$$u^e(s) = -B_D^{out}(s) B_S^T y_p(s), \tag{35}$$

where $B_D^{out}(s)$ matrix captures the dynamics of the edges in the dynamic topology (each column in $B_D^{out}(s)$ describes the dynamic of an edge $e_{ij}(s) = \lambda_{ij}(s)$) and B_S matrix captures

the interconnections between the nodes and the edges in the dynamic network.

If the $B_D^{out}(s)$ matrix has a state space realization (A_e, B_e, C_e, D_e), the dynamic topology model (35) can be written in the time domain as:

$$\begin{aligned} \dot{x}_e(t) &= A_e x_e(t) - B_e B_S^T y_p(t) \\ u^e(t) &= C_e x_e(t) - D_e B_S^T y_p(t), \end{aligned} \tag{36}$$

where $x_e(t) \in \mathbb{R}^{\bar{n}}$ is the state vector of the matrix B_S and A_e, B_e, C_e, D_e are constant matrices with appropriate dimensions.

Combining (30), (31) and (36), the overall representation of the dynamic consensus network with N (LTI) heterogeneous nodes and dynamic edges can be written as:

$$\begin{bmatrix} \dot{x}_p(t) \\ \dot{x}_e(t) \end{bmatrix} = \begin{bmatrix} A_p - B_p D_e B_S^T C_p & B_p C_e \\ -B_e B_S^T C_p & A_e \end{bmatrix} \begin{bmatrix} x_p(t) \\ x_e(t) \end{bmatrix} + \begin{bmatrix} B_p \\ 0 \end{bmatrix} u^{in}(t),$$
$$y(t) = \begin{bmatrix} C_p & 0 \end{bmatrix} \begin{bmatrix} x_p(t) \\ x_e(t) \end{bmatrix}. \tag{37}$$

Define $\widetilde{D_e} = D_e B_S^T$ and $\widetilde{B_e} = B_e B_S^T$, the system (37) can be rewritten as

$$\begin{bmatrix} \dot{x}_p(t) \\ \dot{x}_e(t) \end{bmatrix} = \begin{bmatrix} A_p - B_p \widetilde{D_e} C_p & B_p C_e \\ -\widetilde{B_e} C_p & A_e \end{bmatrix} \begin{bmatrix} x_p(t) \\ x_e(t) \end{bmatrix} + \begin{bmatrix} B_p \\ 0 \end{bmatrix} u^{in}(t),$$
$$y(t) = \begin{bmatrix} C_p & 0 \end{bmatrix} \begin{bmatrix} x_p(t) \\ x_e(t) \end{bmatrix}. \tag{38}$$

From the overall representation of the dynamic network (38), the kernel representation of this system can be expressed by:

$$\begin{bmatrix} sI_n - A_p + B_p \widetilde{D_e} C_p & -B_p C_e & -B_p \\ \widetilde{B_e} C_p & sI_{\bar{n}} - A_e & 0 \end{bmatrix} \begin{bmatrix} x_p(t) \\ x_e(t) \\ u^{in}(t) \end{bmatrix} = 0, \tag{39}$$

Using the above kernel representation (39), the following theorem is a direct result of applying Theorem 3.1 above:

Theorem 4.1: The dynamic consensus network (37) is controllable if and only if the rank of $\begin{bmatrix} sI_n - A_p + B_p \widetilde{D_e} C_p & -B_p C_e & -B_p \\ \widetilde{B_e} C_p & sI_{\bar{n}} - A_e & 0 \end{bmatrix}$ is the same for all $s \in \mathbb{C}$.

A proof of this theorem can be done directly. Alternately, we can use Theorem 9.4 in [13], which states that for two systems S_1 and S_2 connected in feedback configuration shown in Fig. 5, the closed loop system is controllable if S_1, S_2, and the series connection $S_1 S_2$ are each controllable. The later condition requires that no pole of S_2 is a zero of S_1. Using this result we can state the following:

Theorem 4.2: The dynamic network described by (37) is controllable if each of the following is true:

1) The system of nodes defined by (30) is controllable.
2) The system of edges defined by (36) is controllable.
3) No pole of the edge system (36) is a transmission zero [1] of the node system (30).

Proof:

From Theorem 4.1, we conclude that the system (37) is controllable if and only if the $rank(Q(s))$ is the same for all $s \in \mathbb{C}$, where

$$Q(s) = \begin{bmatrix} sI_n - A_p + B_p \widetilde{D_e} C_p & -B_p C_e & -B_p \\ \widetilde{B_e} C_p & sI_{\bar{n}} - A_e & 0 \end{bmatrix}. \tag{40}$$

[1] A transmission zero for $\dot{x} = Ax + Bu$, $y = Cx$ is a value $\lambda \in \mathbb{C}$ such that $\begin{bmatrix} \lambda I_n - A & B \\ C & 0 \end{bmatrix}$ loses rank. Transmission zeros correspond to terms $e^{\lambda t}$ that will not appear in the output when they are in the input.

If we do some column operations on the matrix $Q(s)$ ($Col_1 = Col_1 + Col_3 \widetilde{D_e} C_p$, $Col_2 = Col_2 - Col_3 C_e$), and replacing the columns 2 and 3, we get

$$\bar{Q}(s) = \begin{bmatrix} sI_n - A_p & -B_p & 0 \\ \widetilde{B_e} C_p & 0 & sI_{\bar{n}} - A_e \end{bmatrix}. \tag{41}$$

Note that $\bar{Q}(s)$ is column equivalent to $Q(s)$ and then $rank\bar{Q}(s) = rank(Q(s))$, for all s. The theorem can be proven if we can show that the $rank(\bar{Q}(s)) = n + \bar{n}$ for the following two cases:

- For $s \notin \lambda(A_e)$, condition 1. implies that $rank(sI_n - A_p, B_p) = n$, for all s. For $s \notin \lambda(A_e)$, $sI_{\bar{n}} - A_e$ is nonsingular. Then from the matrix structure of $\bar{Q}(s)$, $rank(\bar{Q}(s)) = n + \bar{n}$ and then $rank(Q(s))$ is the same for all s. Thus the system is controllable.
- For $s \in \lambda(A_e)$: We can rewrite $\bar{Q}(s)$ in (41) as

$$\bar{Q}(s) = \bar{Q}_1(s) \bar{Q}_2(s), \tag{42}$$

where

$$\bar{Q}_1(s) = \begin{bmatrix} I_n & 0 & 0 \\ 0 & \widetilde{B_e} & sI_{\bar{n}} - A_e \end{bmatrix},$$
$$\bar{Q}_2(s) = \begin{bmatrix} sI_n - A_p & -B_p & 0 \\ C_p & 0 & 0 \\ 0 & 0 & I_{\bar{n}} \end{bmatrix}. \tag{43}$$

Form (43), condition 2. in this theorem implies that the $rank(sI_{\bar{n}} - A_e, \widetilde{B_e}) = \bar{n}$, for all s and then $rank(\bar{Q}_1(s)) = n + \bar{n}$, for all s. For $s \in \lambda(A_e)$, the third condition implies that the $rank(\bar{Q}_2(s)) = n + p + \bar{n}$, for all $s \in \lambda(A_e)$, where $p = \sum_i p_i$ is the dimension of the output vector $y_p(t)$ of the system.

Applying the Sylvester's rank inequality (see [14] and [15]) in the equation (42), we get

$$n + \bar{n} \geq rank(\bar{Q}(s)) \geq rank(\bar{Q}_1(s)) + rank(\bar{Q}_2(s)) - d, \tag{44}$$

where $d = n + p + \bar{n}$ is the number of columns in the matrix $\bar{Q}_1(s)$.

From (44) and (43), we can rewrite the last inequality as

$$n + \bar{n} \geq rank(\bar{Q}(s)) \geq (n + \bar{n}) + (n + p + \bar{n}) - (n + p + \bar{n}). \tag{45}$$

From the above equation we can conclude that the $rank(\bar{Q}(s)) = n + \bar{n}$, for all $s \in \lambda(A_e)$ and then the rank of the matrix $Q(s)$ is the same for all s and hence the system is controllable. ■

Note that Theorem 4.1 and Theorem 4.2 are equivalent. The controllability conditions introduced in Theorem 4.2 are interesting because it concludes that in addition to the controllability of the nodes and edges, controllability requires the poles of the edges not match the transmission zeros of the nodes. Furthermore, the third condition can be investigated using only the region of the eigenvalues of the system matrix A_e. This makes controllability analysis easier because it can be investigated using only the bounds of $\lambda_i(A_e)$ or by simply checking the poles and zeros of the nodes and edges dynamics.

V. Conclusion

This paper presented a framework for dynamic graphs and dynamic consensus networks. This framework introduced the idea of dynamic degree, adjacency, incident, and Laplacian matrices in a way that naturally extends these concepts from the static case. From this, equivalent concepts for dynamic interconnection matrices and dynamic consensus networks were defined. We used the established aspects and properties of the defined dynamic graph theory in conjunction with the behavioral approach to develop a controllability-analysis methodology for dynamic networks. Additionally, the application of the dynamic interconnection topology in analyzing the interconnections of nodes for the purpose of establishing controllability analysis was demonstrated. The behavioral approach was described,

followed by a discussion of its use in controllability of the two-rooms example. This led to the presentation of the resulting general usage of dynamic graph theory in conjunction with the behavioral approach for controllability-analysis of dynamic consensus networks.

REFERENCES

[1] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on automatic control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[2] K. L. Moore, T. L. Vincent, F. Lashhab, and C. Liu, "Dynamic consensus networks with application to the analysis of building thermal processes," *Proceedings of the 18th IFAC World Congress*, vol. 18, no. 1, pp. 3078–3083, 2011.

[3] K.-K. Oh, F. Lashhab, K. L. Moore, T. L. Vincent, and H.-S. Ahn, "Consensus of positive real systems cascaded with a single integrator," *International Journal of Robust and Nonlinear Control*, vol. 25, no. 3, pp. 418–429, 2015.

[4] X. Liu, H. Lin, and B. Chen, "A graph-theoretic characterization of structural controllability for multi-agent system with switching topology," in *Proceedings of the 48th IEEE Conference on Decision and Control, held jointly with the 28th Chinese Control Conference*, pp. 7012–7017, IEEE, 2009.

[5] H. Tanner, "On the controllability of nearest neighbor interconnections," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, vol. 3, pp. 2467–2472, IEEE, 2004.

[6] M. Ji, A. Muhammad, and M. Egerstedt, "Leader-based multi-agent coordination: Controllability and optimal control," in *Proceedings of 2006 American Control Conference*, pp. 1358–1363, IEEE, 2006.

[7] A. Rahmani, M. Ji, M. Mesbahi, and M. Egerstedt, "Controllability of multi-agent systems from a graph-theoretic perspective," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 162–186, 2009.

[8] J. Willems, "The behavioral approach to open and interconnected systems," *IEEE Control Systems Magazine*, vol. 27, no. 6, pp. 46–99, 2007.

[9] J. Willems, "On interconnections, control, and feedback," *IEEE Transactions on Automatic Control*, vol. 42, no. 3, pp. 326–339, 1997.

[10] J. Willems and Y. Yamamoto, "Linear differential behaviors described by rational symbols," in *Proceedings of the 17th IFAC World Congress, Seoul, Korea*, pp. 12266–12272, 2008.

[11] J. Polderman and J. Willems, *Introduction to Mathematical Systems Theory: a Behavioral Approach*. Springer Verlag, 1998.

[12] J. Willems, "The Behavioral Approach to Open and Interconnected Systems Modeling by tearing, zooming, and linking," *IEEE Control Systems Magazine*, vol. 27, no. 6, pp. 46–99, 2007.

[13] C. Chen, *Linear System Theory and Design*. Oxford University Press, Inc., 1998.

[14] G. Matsaglia and G. Styan, "Equalities and inequalities for ranks of matrices†," *Linear and Multilinear Algebra*, vol. 2, no. 3, pp. 269–292, 1974.

[15] A. Ramakrishna and N. Viswanadham, "Decentralized control of interconnected dynamical systems," *IEEE Transactions on Automatic Control*, vol. 27, no. 1, pp. 159–164, 1982.

A metaheuristic approach to optimize maintenance productivity of Indian railway coach and bogie

Sanghamitra Poddar[1], Subhash Chandra Panja[2], Sankar Narayan Patra[2], Debanjan Ghosh[1], Debajyoti Saha[1]

1. Institute of Engineering & Management, Salt Lake, Kolkata-700091
2. Jadavpur University, Jadavpur, Kolkata-700032

Abstract- Indian Railways being one of the largest functioning units of India employs a huge amount of manual unorganized labor. Maintenance of railway coaches and bogies is provided from part of this labor. It has always been a great challenge to find an optimal operating scenario for maximized productivity. In this work, we aimed to model and maximize maintenance productivity of Indian Railway coach and bogie . Taking into consideration the volatility of the circumstances and uncorrelated behavior of different human factors affecting productivity, to obtain the best scenario five of the most critical human factors were identified that are expected to affect the maintenance of the coaches and bogies most. Cobb-Douglas production function was used to model maintenance productivity with the chosen human factors. The main goal after designing the model for productivity is to maximize it. To get the optimal productivity Differential Evolution algorithm is used. Finally, the Cuckoo search algorithm is applied to validate the results.

***Keywords*- Railways, maintenance, human factors, productivity, differential evolution, statistical, cuckoo search.**

I. INTRODUCTION

Indian Railways(IR) is the fourth-largest railway network in the world by size and also the world's eighth-largest employer as of March 2017 [1]. Railways play an important role in stimulating the industrial and agricultural demands of the nation. Among different subsystems of IR rolling stock is considered to be the most important as it carries passengers and goods. The rolling stock of IR consisting of locomotives, goods wagons, and passenger coaches, serves 8.26 billion passengers and 16 billion tons of freight per year [2]. Supporting such a multidimensional and extremely important role in a seamless manner, the coaches and wagons suffer intense stress due to climate, the topography of the service area, the tremendous loading of human (passenger services), and goods (freight services), average running time and average distance covered. The maintenance workforce lies at the heart of the system, as having regular maintenance can help us get uninterrupted continued service. Despite a few distracting scenarios like delays and accidents, the primary target of Indian Railways would be to maximize the productivity of the maintenance workforce. A study by Poddar et al. (2017) underlined the key human factors that affect maintenance work in Indian Railways [3]. Due to the dense network of rail transport in India and the high frequency of operations, timely supervised maintenance is an utmost need. Data standardization and Principal Component Analysis (PCA) were used on the list of thirty-six identified human factors. The most important human factors were identified using Factor analysis. In this context, the primary objective of this research work is to present a working model of productivity as a function of various key parameters and maximize it using the Differential Evolution algorithm.

The paper is divided into four main sections. The basic background of the problem and a brief account of the available literature is discussed in section II. After a brief discussion on the key parameters of the Differential Evolution (DE) algorithm in section III, the research framework and workflow are outlined in section IV. Results are discussed in section V with appropriate conclusions being drawn finally in section VI. These sections are discussed in detail.

II. PROBLEM FORMULATION

A. Dependency on human factors:

Productivity being directly linked to the human workforce, it is one of the most important and well studied components of the industry. Endsley (1995) discussed the role of attention and working memory in dynamic human decision making [4]. Wilson and Rutherford (1989) worked on psychological impact of change in human factors [5]. In their study on Human Robot Interaction, Hancock et al. (2011)[6] found robot performance and attributes to be the largest contributors in increasing trust in Human Robot Interaction. Hendrick (1991) and Krüger et al. (2009) also presented similar outcomes [7][8]. Keller *et. al.* discussed the necessity of maximum eight hour single shifts to avoid

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

fatigue, health issues leading to reduction in productivity. This observation was supported by Cotrim et al. (2017)'s study showing rest quality being an important factor [10]. From Indian perspective, it is very important to conceive a proper work shift pattern for the maintenance formalities due to the huge number of people availing the services [11, 12]. The human factor also drastically affects the safety issues in railways. Baldry and Ellison (2006) worked extensively to identify instances of major incidents involving rail employees [13].

B. Significance of the Productivity Analysis:

Human factors is one of the most important parameters which is being analysed in various industries, as productivity is greatly dependable on it[14-16-DE]. In Order to increase the productivity mutual dependencies of some key factors are sought out. Usually, the increase in productivity is due to the increasing favourability of these factors.

C. The Cobb-Douglas function:

Productivity model is a measurement method which is used in practice as a ratio of value added output and different inputs like material, capital, land, man and machine.

Keeping this in mind the present work deals with estimating the total factor productivity (TFP) of maintenance workforce for railway coaches and bogies.

On the basis of parameters like human machine interaction, time pressure, situational awareness, safety culture and shift pattern, the well-known Cobb-Douglas production function outlines its input-output relationship [17-18]. It allows the study on the impact of humane factors upon above mentioned factors onto input labor parameters of the function.

Yield parameters within the Cobb-Douglas production functions are communicated in exponential shape. This complicates the ensuing analysis and application. To simplify the reliance P H Douglas (1976) [19-DE] tested with diverse viewpoints of the productivity work and eventually communicated the function in terms of direct homogeneous functions utilizing logarithmic changes. This transformation simplified the application of statistics (Goldberger) [20-DE].

This production function is in this way utilized to set up a working connection between the productivity of Indian Railroads coach and bogie support and the human components administering the performance of the workforce. The productivity function is then maximized utilizing Differential Evolution optimization techniques.

D. Productivity Function Optimization:

In previous research Least Square method (LSM) and Particle Swarm optimization (PSO) techniques have been implemented to maximize the productivity, by the authors.

E. Application of Differential Evolution Algorithm:

First introduced in 1995 by K V Price and R M Storn [21], Differential Evolution Algorithms have been popular in various optimization problems. In this work the authors have applied the Differential Evolution (DE) technique on the Cobb-Douglas function to maximize productivity. A brief description of Differential Evolution algorithm used along with optimized parameter values is discussed in consequent segments.

III. BASIC DIFFERENTIAL EVOLUTION (DE) FORMULATION

A. Initialization and choice of starting population:

Like almost all evolutionary algorithms, DE starts from an initial population containing its constituent members distributed over a bounding box constrained by the limits on the parameters. The upper and lower bounds of the parameters are either intuitively chosen from physical understanding or are dynamically updated across the optimization routine.

After the initial bounding box is created from the upper ($b_L b_U$) and lower limits ($b_L b_L$) on the parameters for optimization, the next step is to distribute the entire population inside that bounding box. Uniform or non-uniform distributions can be used for this purpose. The choice of the distribution to be used depends on the knowledge of the location of the global optimum.

If the global optimum point is completely unknown and can be anywhere inside the parameter region a uniform distribution can be used.

$$x_{j,i,0} = b_{jL} + rand_j(1, D)(b_{jU} - b_{jL}) \quad (1)$$

Here 'D' is the number of parameters in the optimization problem. On the other hand if there is some information regarding the location of the optimum point, a Gaussian distribution can be used to emphasize the region containing the optimum.

B. Differential Mutation and advantages

The process of mutation is completely unique in the case of the DE algorithm and sets it apart from other available evolutionary algorithms. While GA uses a bit flipping technique, DE uses differential mutation wherein a scaled difference of two randomly chosen members is added to a third randomly chosen member. The end result is a mutated vector.

$$v_{i,g} = x_{r_0,g} + F(x_{r_1,g} - x_{r_2,g}) \quad (2)$$

There are four major parameters in case of differential mutation described by the above mentioned equation. The base vector index (r_0), the difference vector indices (r_1 and r_2) and the mutation scale factor (F).
The standard procedure of randomly choosing the vector indices (also called the Roulette wheel selection) leads to selection duplication. This error is eliminated using Stochastic Universal Sampling where all the vector indices are chosen randomly but simultaneously at one go.
The two vectors $r_1 r_1$ and r_2 for differencing are selected in the similar fashion as was done in case of $r_0 r_0$, but with different and individual permuted arrays. In this way it is ensured that $i \neq r_0 \neq r_1 \neq r_2$

In any evolutionary algorithm, mutation always brings diversity and consequently an option for better performing members in the population. The factor that controls the amount of mutation is the mutation scale factor F as defined in (1). Increasing F would indeed increase diversity in the population leading to a better performance but it has been seen that with F>1 the solutions are both time consuming and less reliable than with F<1. Also when F=1, identical mutated vectors are created which degrades the performance of the DE technique.

C. Crossover Formulation:

Crossover is a process in which a trial vector is created using parameter value selection from the mutant and the target vector. There are various types of crossover technique namely one point crossover, N point crossover and exponential crossover. One parameter is initially chosen at random and copied from the mutant to the corresponding trial parameter. The source of subsequent trial parameters is determined by comparing a cross-over factor (Cr) to a uniformly distributed random number between 0 and 1 that is generated (D-1) times i.e. anew for each parameter. This crossover factor controls the amount of part/parameter exchange between the parent and the subsequent generation thereby controlling variability in characteristics.

D. Selection and termination :

After successful creation of N_p trial vectors through mutation and crossover, the next step is selection of population members for the next generation. The trial vectors and target vectors compete against each other for a place in the next generation. Usually selection is based on objective function value and depending on the optimization problem; either the vector with highest objective function value or the one with lowest objective function value survives and enters the next generation population.

There can be various types of termination criteria in any optimization technique. Some of them are being discussed here. An absolute value of the productivity function can be set as the barrier. Alternatively a differential convergence based adaptive termination criteria can be set where a predefined percentage of change of the productivity function in successive iterations can be sought. For a failsafe, a maximum number of iterations/generations are kept such that the optimization procedure is not stuck in an infinite non-evolving loop.

IV. RESEARCH FRAMEWORK

A. Overview:

It has been built up in the last segments that the maintenance of the railway bogies and coaches is highly dependent on the performance of the employees. A rudimentary survey on the railway department employees has been done to estimate the human factors and their influence and dependency on productivity. The modelling of the productivity function is done based on the collected statistical data.

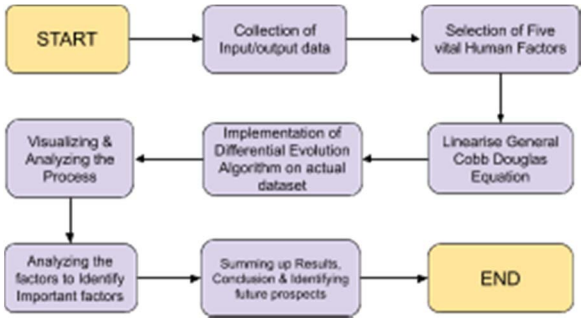


Fig. 1. Flowchart of the research work

B. Dataset Overview:

In this work, data were collected mainly from interactions with various railway operation and maintenance staff. Initially the documentation started with interviewing 250 people at different levels of railway operation and maintenance to explore the IR maintenance system and to identify associated human factors. An open ended prototype questionnaire was formulated. This allowed freedom for both the interviewer and the interviewee encouraging supplementary questions (interviewer) and wilful response (interviewee). Frequent visits to the maintenance work shed in the context of historical railway operation and maintenance data available in the public domain led to the identification of different model parameters.

The identification of human factors was influenced by:

a) registers/files/documents/data sheets related to maintenance and operation of IR

b) direct observations of maintenance activities and operations

c) discussion with concerned operators and staff engaged in the operation of maintenance of railway coach and bogie.

To obtain a comprehensive dataset, these factors were extensively studied with repeated interactions with working personnel ranging from administrators to ground staff. Qualitative factors like expectation, interpretation, perception, situational awareness, human-machine interface, and vigilance are mapped to quantifiable factors by marking them in a scale of '1 to 10' based on its importance as felt by the different interviewees. The marking system correlated human factors based on their importance with respect to maintenance of coach and bogie system.

C. Building Model :

The model developed form the dataset thus collected deals with the role of human factors on productivity using standard Cobb-Douglas production function. The most basic form of the Cobb-Douglas function for a single good with two input factors is:

$$Y=A\,L^{\beta}\,K^{\alpha} \quad (3)$$

where:
Y = total production
L = labour input
K = capital input
A = total factor productivity
α and β are the output elasticity of capital and labour, respectively. Output elasticity measures the responsiveness of output to a change in levels of either labour or capital used in production.
To handle multiple inputs, the Generalized Cobb-Douglas (CD) production function is introduced which can handle multiple inputs. It is represented as

$$f(x) = A \prod_{i=1}^{L} x_i^{\lambda_i} \quad x = (x_1, \ldots x_L) \quad (4)$$

where:
1) A is an efficiency parameter
2) L is the total number of goods
3) $x_1, ..., x_L$ are the (non-negative) quantities of good consumed, produced, etc.
4) λ_i is an elasticity parameter for good i

The input factors obtained from the dataset are mapped into the multiple inputs of the Cobb-Douglas production function.

D. Analysis:

Let X denote the m×n matrix representation of the available input dataset from statistical observations as discussed earlier. Here m denotes the number of observations made corresponding to each of the n human factors. We have m=250 and n=5. The i^{th} row of the input dataset Xi is thus expressed as
$Xi = [x_{iA}\, x_{iB}\, x_{iC}\, x_{iD}\, x_{iE}]$

Let Y be the m×1 matrix consisting of the output productivity values for each of the m=250 observations.

Assuming a basic production function estimate, productivity output (Y) can be denoted as a function of inputs of Human Machine Interface (A), Situational Awareness (B), Time Pressure (C), Shift Pattern (D), Safety Culture (E). We have:

$$Y=f(A,B,C,D,E) \quad (5)$$

Assuming Cobb-Douglas estimate:

$$Y=A^aB^bC^cD^dE^e \quad (6)$$

Contributions of each of the factors are controlled through the corresponding elasticity parameters (a, b, c, d and e). Taking logarithm of equation (2) we obtain equation (7):

$$\log(Y)= a\log(A) + b\log(B) + c\log(C) + d\log(D) + e\log(E) \quad (7)$$

Thus, our assumption of production function reduces to a linear estimate with parameters a, b, c, d, e.

V. RESULTS AND DISCUSSION

The DE algorithm discussed in section III is applied on (7) to maximize the productivity function log(Y) (and consequently Y). DE parameters and performances are discussed next.
In the context of discussion in section III, we have chosen flexible upper and lower bounds on the elasticity parameters of (7). Depending on overall convergence, the bounds are decided. A uniformly distributed population space of 100 members is chosen to initialize DE. In case of differential mutation a universal stochastic sampling is chosen to obtain the base vector as well as the differencing vectors. A random permutation matrix is introduced to avoid duplication in vector selection. The mutation control factor of 0.85 is chosen to obtain a higher degree of mutation in the solution space. Post mutation, a uniform crossover scheme is adopted with a crossover probability of 0.2. Consequently a fairly diverse population space is generated after each generation. This had been a single objective optimization with a goal to maximize the productivity (Y). The termination criteria are set as a 5% change in subsequent iterations or a normalized productivity value of 0.99 or a maximum iteration of 1000 whichever is achieved earlier. It is observed that the optimization algorithm runs for 200 iterations before reaching a differential accuracy of 5% and the termination criteria is hit.

A. Optimization of Parameter using DE Algorithm:

The optimal elasticity coefficient values obtained by Differential Evolution Algorithm on our available dataset.Values of each coefficient is between 0 and 1. Keeping other inputs as constant we are measuring the change of output productivity for each and every individual coefficient.The effect of individual input factors on overall productivity output is clearly visible here. It is observed Human Machine Interface with the maximum coefficient of 0.567 is the most significant human factor contributing to maintenance service which is followed by Safety Culture (0.342), Shift Pattern(0.208), Time Pressure (0.188) and Situational Awareness (0.165). In order to validate the mathematical model and to ensure that randomness of Differential Algorithm isn't led to a faulty result,a nature inspired searching algorithm (Cuckoo Search) is applied to the same dataset for similar iteration and same other parameter constraints.The result obtained using both the algorithm is tabulated in Table 1.

TABLE 1: ELASTICITY COEFFICIENT OF INDIVIDUAL FACTORS

Human Factor	Elasticity Coefficient	Value obtained (by DE)	Value obtained (by Cuckoo Search)
Human Machine Interface	a	0.567	0.553
Situational Awareness	b	0.165	0.173
Time Pressure	c	0.188	0.187
Shift Pattern	d	0.208	0.195
Safety Culture	e	0.342	0.334

It is observed that with both the algorithms values obtained are very close to each other and both point that Human Machine Interface is the most important factor affecting the productivity of the maintenance workforce.

Figures 2.1-2.10 show the convergence plot of Differential Evolution(DE) along various axes.

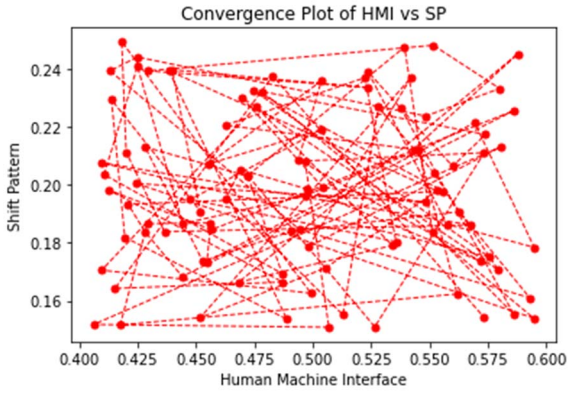


Fig. 2.1

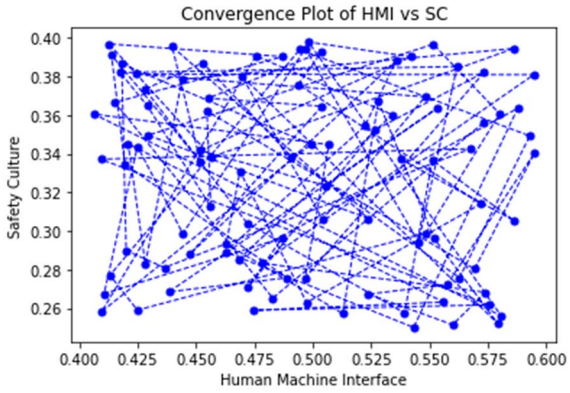


Fig. 2.2

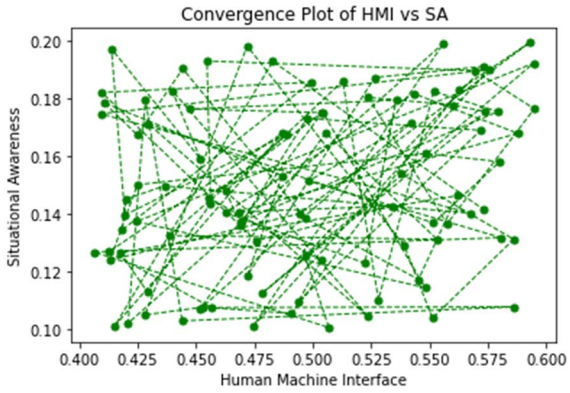


Fig. 2.3

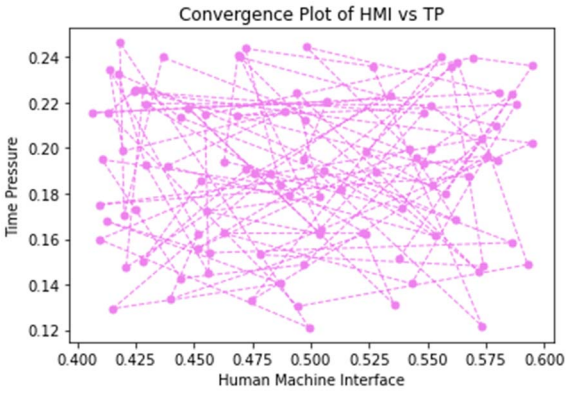


Fig. 2.4

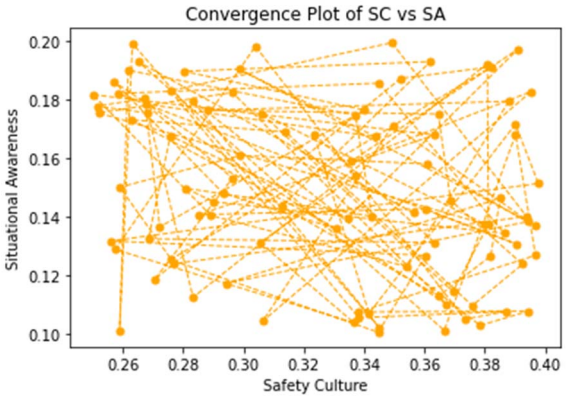


Fig. 2.5

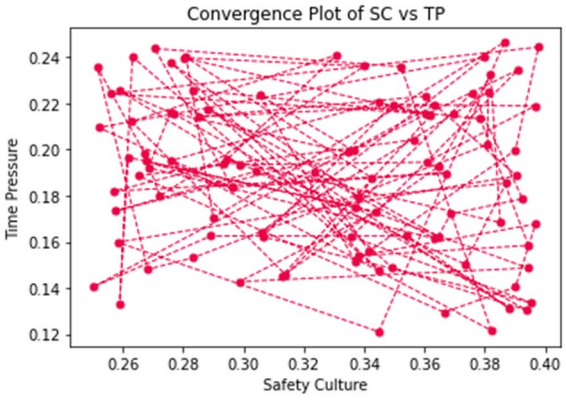


Fig. 2.6

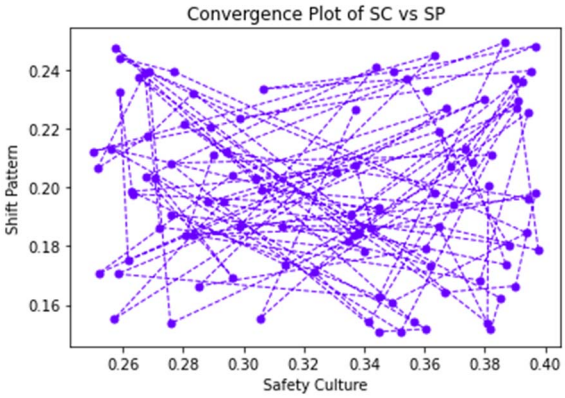


Fig. 2.7

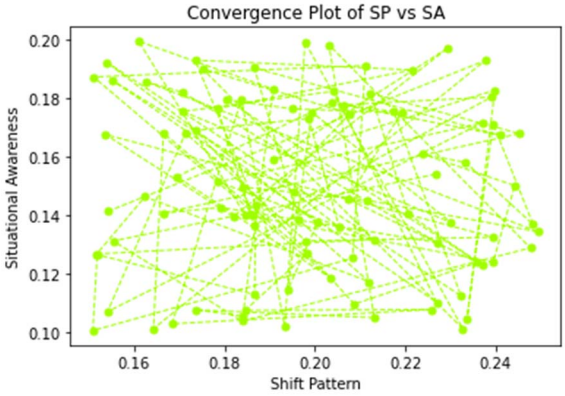


Fig. 2.8

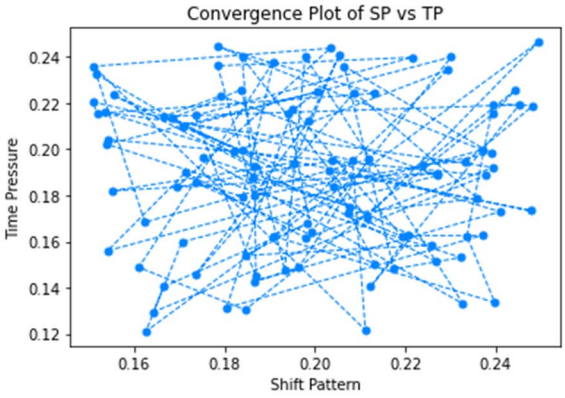


Fig. 2.9

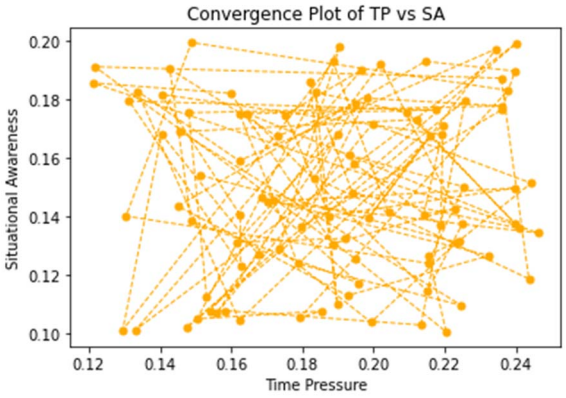


Fig. 2.10

Figure 2.1-2.10 show convergence plot of all Five Parameters

Figure 3.1-3.5 shows the correlation between the predicted and actual values of all the selected five parameters. The few extreme points in each plot are considered as the outliers, observations show that the predicted values by the DE algorithm lies close to the actual values. The actual observations in productivity values are represented in blue dots, and the predicted productivity output are represented in orange stars.

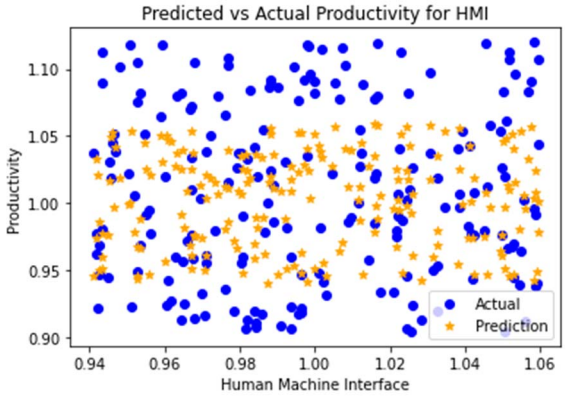


Fig. 3.1

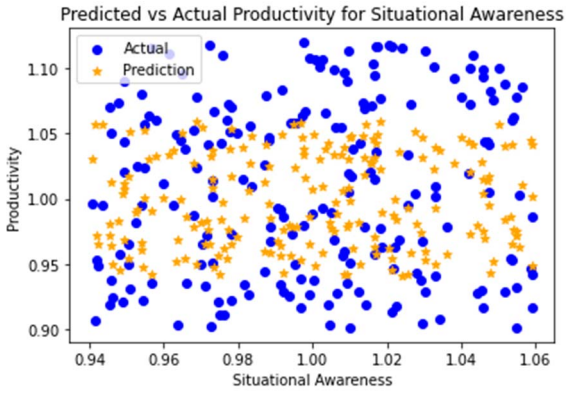


Fig. 3.2

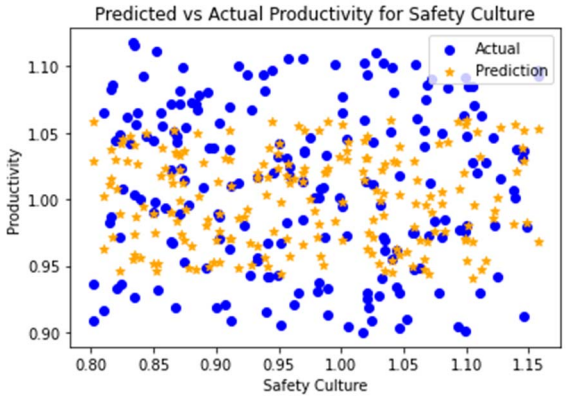


Fig. 3.3

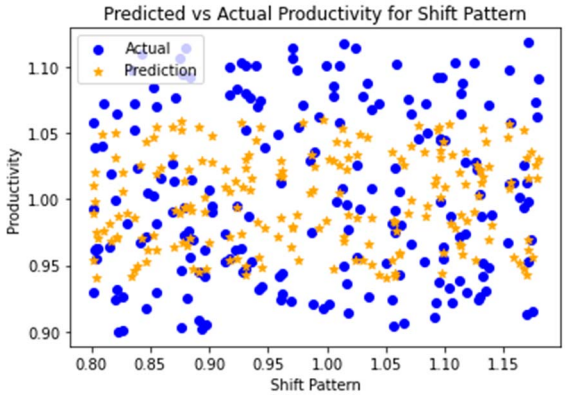


Fig. 3.4

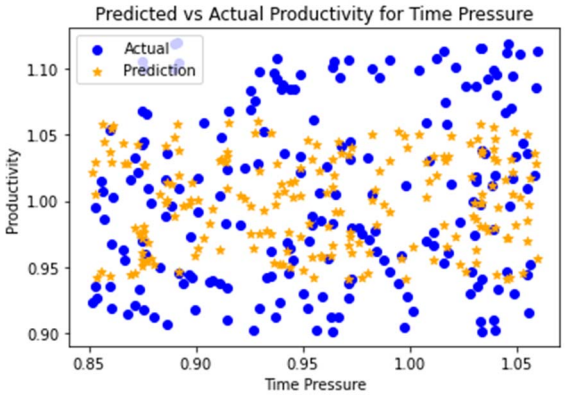


Fig. 3.5

B. Model training and Loss Chart:

This research focussed on training of the DE algorithm to achieve certain parameter values in order to maximise the productivity value. It is observed that the initial loss value was quite large and which gradually reduced over the entire training procedure of 200 epochs. Figure 4, depicts the loss curve during the complete training process.

Loss vs Epochs

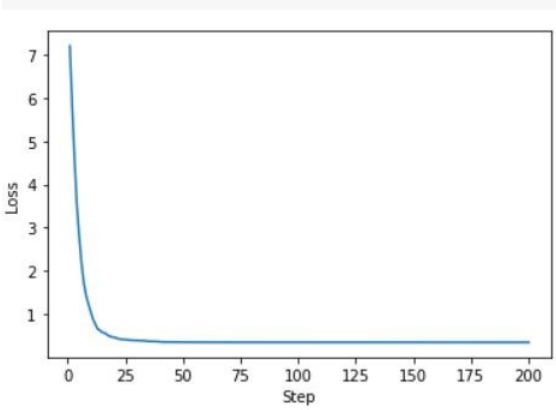


Fig.4

C. Output Productivity Maximization:

The maximum productivity value is computed to be 9.987 taking the factor scores as 10. The maximum attainable range is taken to be within 5% of this value.

Table-2 depicts the range of each individual input parameter for output to fall in the specified range.

Table-2 Individual Factor for 5% Output Maximization

Human Factor	Minimum value	Maximum value
Human Machine Interface	7	10
Time Pressure	6	10
Situational Awareness	8	10
Shift Pattern	2	10
Safety Culture	2	10

The plot of these individual five parameters with productivity show the range of the input parameters to obtain the maximum maintenance productivity.

Figure 5.1-5.5, depicts the spread of the individual factors(Human Machine Interface, Time Pressure, Situational Awareness, Shift Pattern and Safety Culture). In each plot, Y-axis represents the range of output values considered for maximization. X-axis represents the values taken by parameters for respective outputs. It shows a range of values for each parameter for which the output tends maximum productivity, thereby justifying each of the aforementioned results in table-2..

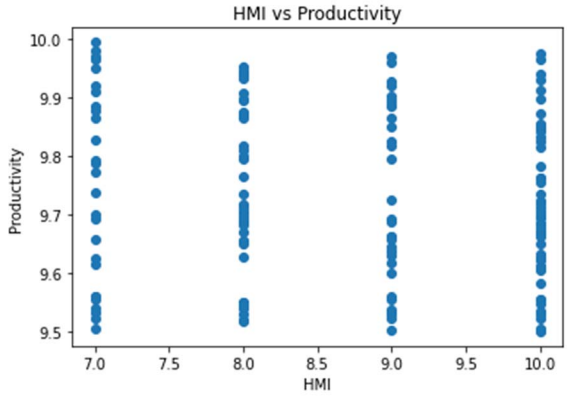


Fig. 5.1

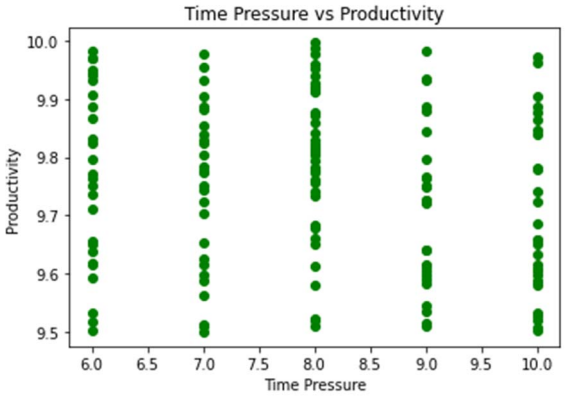


Fig. 5.2

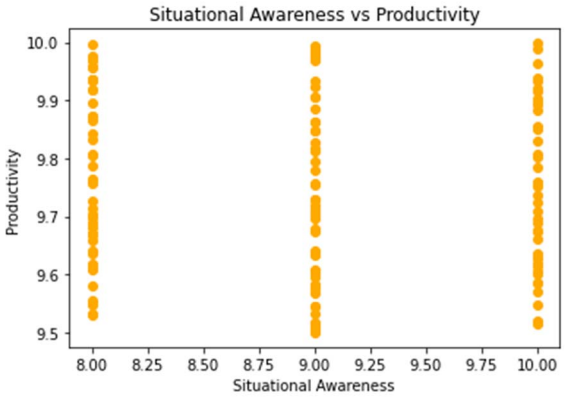


Fig. 5.3

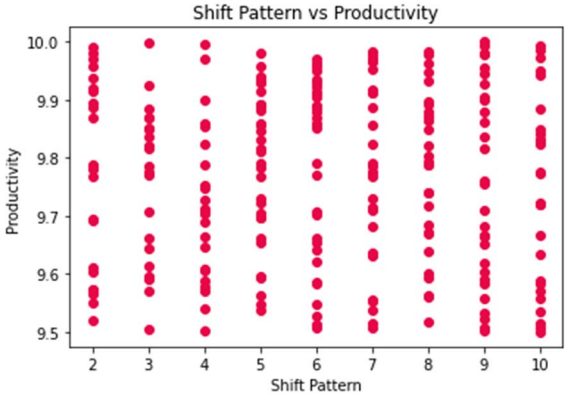


Fig. 5.4

Fig. 5.5

Figure 5.1-5.5 Variation of Productivity with all the Five input factors.

VI. CONCLUSION

In this work, it was mathematically established which are the important human factors affecting the overall performance of maintenance of coach and bogie of Indian Railways system. Five out of thirty-six human factors were chosen from the study conducted by Poddar et. al. for productivity analysis, a production function was derived using surveyed datasets based on the standard Cobb-Douglas function. This production function was then maximized using the Differential Evolution algorithm to obtain an optimal operating condition in terms of the elasticity factors of the five chosen human factors. It was observed that the results obtained indicate that out of the five human factors chosen for productivity management, the optimal scenario is achieved when maximum weight is assigned to the human-machine interface factor. Subsequent orders of decreasing weights/importance are found to be Safety Culture(E), Shift Pattern (D), Time Pressure(C) and Situational Awareness(B). DE uses a manageable resource in terms of computation time and computing power requirement to obtain these results. Previously reported optimization efforts using Cuckoo search method indicate and support the findings of this work.

VII. Scopes of future work

The model can be implemented in various places which desire to achieve maximum productivity of its workmen. The method provides a valid assessment of the current workflow process and provides results for the most productive way to perform maintenance operations. In future, this concept can be implemented to develop an application software that can perform quick assessment of overall productivity.

REFERENCES

[1] Ministry of Railway (2018, February). Indian Railways Statistical Publications 2016-17: *Statistical Summary – Indian Railways*

[2] Ministry of Railway (2018, February). *Indian Railways Budget Documents 2018-19*.

[3] Poddar, S., Panja, S.C., Gangopadhyaya, M. (2017). "Human Factor Analysis for Railway Coach and Bogie Maintenance using PCA & Factor Analysis."

[4] Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human factors*, *37*(1), 32-64.

[5] Wilson, J. R., & Rutherford, A. (1989). Mental models: Theory and application in human factors. *Human Factors*, *31*(6), 617-634.

[6] Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, *53*(5), 517-527.

[7] Hendrick, H. W. (1991). Ergonomics in organizational design and management. *Ergonomics*, *34*(6), 743-756.

[8] Krüger, J., Lien, T. K., & Verl, A. (2009). "Cooperation of human and machines in assembly lines." *CIRP Annals-Manufacturing Technology*, *58*(2), 628-646.

[9] Keller, S. M., Berryman, P., & Lukes, E. (2009). "Effects of extended work shifts and shift work on patient safety, productivity, and employee health" *Aaohn Journal*, *57*(12), 497-504.

[10] Cotrim, T., Carvalhais, J., Neto, C., Teles, J., Noriega, P., & Rebelo, F. (2017). Determinants of sleepiness at work among railway control workers. *Applied ergonomics*, *58*, 293-300.

[11] Ahasan, R. (2002). "Human adaptation to shift work in improving health, safety and productivity–some recommendations." *Work Study*, *51*(1), 9-16.

[12] Jansen, N., Kant, I., van Amelsvoort, L., Nijhuis, F., & van den Brandt, P. (2003). Need for recovery from work: evaluating short-term effects of working hours, patterns and schedules. *Ergonomics*, *46*(7), 664-680.

[13] Baldry, C., & Ellison, J. (2006). Off the rails: factors affecting track worker safety in the rail industry. *Employee Relations*, *28*(3), 255-272.

[14] Lieth, H. (1975). "Modeling the primary productivity of the world.", *Primary productivity of the biosphere* (pp. 237-263*). Springer, Berlin, Heidelberg*.

[15] Coelli, T. (1995). "Estimators and hypothesis tests for a stochastic frontier function: A Monte Carlo analysis." *Journal of productivity analysis*, *6*(3), 247-268.

[16] Arnold, J. M., & Hussinger, K. (2005)"Export behaviour and firm productivity in German manufacturing: a firm-level analysis" *Review of World Economics*, *141*(2), 219-243.

[17] Benos, N., & Karagiannis, S. (2016) "Do education quality and spillovers matter? Evidence on human capital and productivity in Greece", *Economic Modelling*, *54*, 563-573.

[18] Bresnahan, L., Coxhead, I., Foltz, J., & Mogues, T. (2016), "Does Freer Trade Really Lead to Productivity Growth? Evidence from Africa." *World Development*, *86*, 18-29.

[19] Douglas, P. H. (1976). "The Cobb-Douglas production function once again: its history, its testing, and some new empirical values." *Journal of Political Economy*, *84*(5), 903-915.

[20] Goldberger, A. S. (1968) "The interpretation and estimation of Cobb-Douglas functions", *Econometrica: Journal of the Econometric Society*, 464-472.

[21] Kenneth V. Price, Rainer M. Storn, Jouni A. Lampinen, "Differential Evolution - APractical Approach to Global Optimization'', *Springer-Verlag Berlin Heidelberg 2005.*

Multistatic Radar Imaging for Traffic Monitoring

Ozan Dogan
Faculty of Electrical Engineering, Mathematics and Computer Science
Technical University of Delft
Building 36, Mekelweg 4, 2628 CD, Delft, Netherlands
Email: o.dogan@tudelft.nl

***Abstract*—This work presents a novel radar system for vehicle imaging. The proposed system is based on estimation of target parameters like range and velocity and generation of the final radar image. For this purpose, a multistatic Frequency Modulated Continuous Wave (FMCW) Synthetic Aperture Radar (SAR) system is implemented. Both the simulation and on site measurements show that such a radar system has the potential to replace the current very expensive sensors.**

I. Introduction

A considerable number of technologies are currently available for traffic flow monitoring in the literature [1]. These sensors are installed to gather information about the flow of the traffic within a cross section of the road. The major aim is to estimate the parameters like the lane occupancy, number of vehicles per lane, velocity per vehicle, average velocity per lane and the classification of the vehicles by their lengths.

Three sensors are frontiers for extracting these information: The video cameras [2], the inductive loop sensors and LIDARs. The cameras are very effective for imaging the targets however they are light sensitive and demands extensive maintenance due to the dirtiness that makes the cameras blind in compelling weather conditions. Unlike cameras, LIDARs are active devices that are not sensitive to the change of light but still requires maintenance; like the lenses must be frequently cleaned. One of the most efficient and highly accurate sensor is the inductive loop. Nevertheless, just like video cameras and LIDARs, inductive loops require intrusion to the road. The inductive loops shall be installed to the ground and if a failure occurs, the asphalt on the road shall be reconstituted. The camera and the LIDAR shall be installed to the tags that are constructed on the road. If maintenance of the sensor is required, part of the road shall be temporarily closed. In addition, these sensors shall be installed for each lane; results in a very cost inefficient solution.

The Linear Frequency Modulated Continuous Wave (LFMCW) radar is commonly used in many applications including automotive driver-assistance, room-occupancy sensing and industrial automation [3]. There are some efforts for using the radars in traffic monitoring applications and vehicle classification. In [4], micro-Doppler signatures are used for the classification in the traffic intersections. The input of the classifier is only the range of the target. In [5], estimation of range, velocity and acceleration is achieved to extract the range-Doppler image of the scene by installing a front fire radar. However, the spatial (range-azimuth) image of a target is totally different from the range-Doppler image that is obtained in these methods. In order to extract the two dimensional image of the target, the radar may be installed such that the direction of the radar is perpendicular to the direction of the traffic flow; this geometry is termed as side fire. A major advantage of this geometry is that the azimuth profile is not dependent to angular accuracy of the sensor as is the case for the front fire geometry.

In the side fire geometry, the azimuth profile is a function of range and velocity of the target and the antenna pattern of the radar. A reliable length estimation is achievable only in the case that both of these parameters are known. The range estimation is straightforward with the current LFMCW waveforms. However the velocity estimation with the traditional range Doppler method is no longer valid as the relative velocity is quite different from the actual one in the side fire geometry. For this purpose, one way is to apply an Inverse Synthetic Aperture Radar (ISAR) signal processing algorithm for estimating the velocity. The traffic monitoring applications like red light or speed enforcement require a highly accurate and reliable speed measurement that cannot be achieved with the current ISAR techniques. Another method is to increase the number of sensors along the road [6] so the transition duration of the vehicle from one radar to the other can be estimated. By this way, a dual radar can be used for estimating the velocity of the vehicles by using the Displaced Phase Center Antenna (DPCA) techniques.

The SAR principle is a quite mature technique that can be used for imaging vehicles for the traffic monitoring purpose. There is even an effort for satellite SAR imaging of the fixed automotive vehicles [7]. However the moving target imaging is much more complicated.

In this work, a multistatic LFMCW radar system to obtain the ISAR image of vehicles is proposed. By using this system, it is shown that both the range, velocity and range-azimuth image of the target can be extracted in the side fire geometry. In the next section, the problem definition and the imaging algorithm will be explained. Then the parameter estimation will be described. In the fourth section, the algorithm flow will be explained. The fifth section is dedicated to the simulation and on site tests. Last section is the conclusion.

978-1-7281-9615-2/20/$31.00

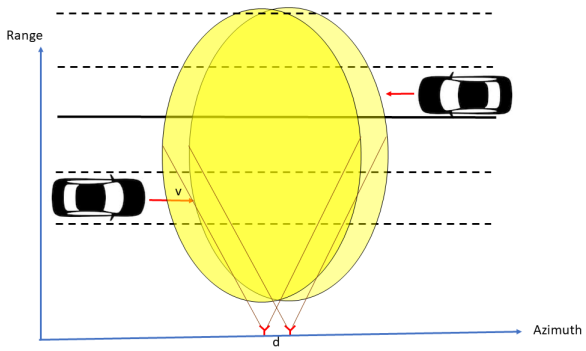


Fig. 1. The traffic monitoring system

II. Problem Definition

The proposed side fire geometry is shown in Figure 1. The radar line of sight is perpendicular to the direction of the traffic flow. There are two radars that have overlapping antenna footprints. The radars are located along the road traffic flow direction with a distance d apart from each other.

The transmitted signal is an LFMCW that can be expressed as follows:

$$s_t(t) = exp(j2\pi(f_c t + \frac{\alpha}{2}t^2) \tag{1}$$

The received signal from a point target at an initial range and azimuth position (r_0, x_0) and an azimuth velocity of v is

$$s_r(t) = a_t(r_0,\eta_0)exp(j2\pi(f_c(t-\frac{2R}{c})+\frac{\alpha}{2}(t-\frac{2R}{c})^2) \\ W_a(\eta,\eta_0,\eta_{dc}) \tag{2}$$

where a_t is the amplitude of the received signal, R is the range that can be expressed as:

$$R = \sqrt{r_0^2 + (vt + v\eta - x_0)^2} \tag{3}$$

where η is the slow time and W_a is the antenna pattern, $\eta_0 = x_0/v$ is the initial time of the target, η_{dc} is the azimuth time of the Doppler centroid. The FMCW radars are mostly homodyne receivers that mixes the transmitted signal and the received one directly in order to decrease the bandwidth of the signal. After mixer and low pass filtering, the IF signal can be written as follows:

$$s_{IF}(t,\eta) = \gamma(r_0,\eta_0)exp(-j2\pi(f_c+\alpha t)\frac{2R}{c})W_a(\eta,\eta_0,\eta_{dc}) \tag{4}$$

where γ represents the amplitude of the received signal after the receiver antenna and all the succeeding amplifiers, attenuators and filters that are assumed to be phase preserving.

In the pulsed SAR systems, the duration of the pulses are mostly very short. So the target movement within the pulse is almost always neglected. This is known as the stop and go approximation that the movement of the target between the receive and transmit is neglected. However, in the FMCW case, this assumption is no more valid [8].

In order to express the impact of the target movement within the chirp, the Fourier transform in the azimuth domain is derived analytically. Let us denote $K_r = \frac{2(f_c+\alpha t)}{c}$ as this term includes solely the fast time. The Fourier transform according to the azimuth time of the signal can be written as follows:

$$S_{IF}(t,f_\eta) = \int s_{IF}(t,\eta)exp(-j2\pi f_\eta \eta)d\eta \tag{5}$$

The principle of stationary phase [9] can be applied in order to derive the Fourier transform of the received signal:

$$s_{IF}(t,f_\eta) = \gamma(r_0,\eta_0)exp(-j2\pi r_0\sqrt{K_r^2 - \frac{f_\eta^2}{v^2}}) \\ exp(-j2\pi(-f_\eta t + f_\eta x_0/v)) \\ W_a(f_\eta,\eta_0,f_{dc}) \tag{6}$$

The first exponential is known to be the Fourier transform of the traditional stop-and-go approach SAR signal, while the second exponential include the movement of the target within the chirp and the initial timing of the target. To develop a SAR imaging scheme, as it is presented in the equation, range and velocity to the target, Doppler centroid frequency and the initial time of the target shall be estimated. The ISAR imaging can be achieved after compensation of the second exponential and Doppler centroid. The remaining term is belonging to the classical ISAR signal and there are many methods for image formation. In this work, a 2D matched filtering is used to reveal the potential of the proposed system.

III. Parameter Estimation

Traditional ISAR imaging algorithms assume a single target within the area of interest. However this is not the case for the traffic monitoring scenario as it is shown in Figure 1. In this case, there may be multiple targets in different ranges with quite different velocities. Hence, an isolation of the signal corresponding to the target is required. Targets have at least three distinctive properties that can be estimated by the radar: Range, velocity and the length. Here, the targets are discriminated in range due to its easiness.

A. Range Estimation

To apply the 2D matched filtering, the estimation of the range is not required. The reference filter can be generated to cover all the possible instrumented ranges. However, this method is very computationally complex. In addition, to estimate the velocity of the target, the range profiles will be already extracted. So a detection algorithm can be used to isolate the targets and dramatically decrease the number of ranges of interest. This is also not only an improvement but also a prerequisite for the traffic monitoring radars that mostly includes very limited signal processing hardware.

After taking the FFT of the signal and the thresholding, the range to the target can be estimated via:

$$f_b = \alpha\frac{2R}{c} \tag{7}$$

where f_b is the beat frequency.

B. Velocity Estimation

The velocity of the target is estimated by the DPCA method with the dual radar system that is presented in Figure 1. In this case, the two azimuth profiles is a shifted version of each other as far as identical radars are assumed. The time difference of the azimuth signal between channels can be expressed as:

$$\Delta t = \frac{d}{v} \tag{8}$$

The velocity is estimated by cross correlating the two channels and computing the shift of the peak. In order to increase the accuracy of the estimation, the distance between the radars must be longer. The upper limit is that the scattering from different azimuth angle must still satisfy a sufficiently high cross correlation. The relation between the accuracy percentage and the distance between the radar, d, can be written as follows:

$$Accuracy(\%) = 100\frac{v_{max}\tau}{d} \tag{9}$$

where v_{max} is the maximum velocity of the vehicle and τ is the duration of the chirp. The accuracy of the estimation can be improved by interpolation [10] or super resolution methods [11].

C. Doppler Centroid Frequency (DCF) Estimation

In the side fire geometry, the radar boresight is aimed to be perpendicular to the road traffic flow direction. Otherwise a shift in Doppler bandwidth occurs and the SAR imaging window moves out of the boresight of the antenna. This may result in poor radiometric resolution for the final SAR image. The radar can be fine tuned so as to satisfy zero DCF. However, the radars may be intentionally installed to some locations like a turnout that the multiple lanes are not satisfying this condition. In order to insert more flexibility to the both installation procedure and location of the radar, a squinted geometry is not only demanded but also unavoidable. The DCF and its relation with the azimuth time, η_{DC} that is mentioned in Equation 4 can be expressed mathematically as:

$$f_{dc} = \frac{2v^2}{\lambda R}\eta_{dc} \tag{10}$$

In the proposed system, the DCF is variant for each target as the velocity of each target is different; something specific to ISAR and not SAR that the DCF can be estimated by using the whole data. For the proposed system, the range profiles are extracted and given as input to the DCF estimation step. There are various methods for estimation of the DCF. In this work, the DCF is estimated by splitting the power spectrum into two and finding the frequency where the energy is as close as possible in both sides as it is described in [12].

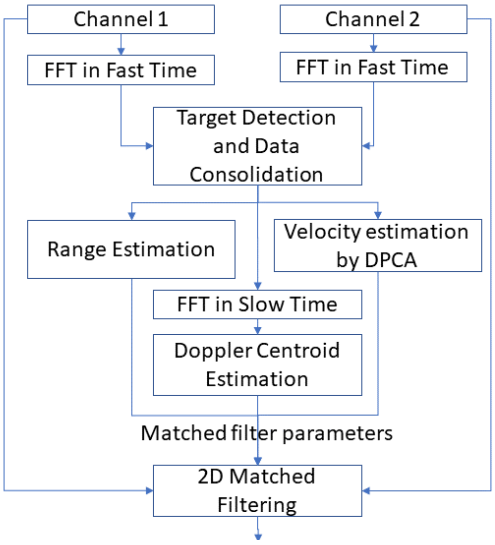


Fig. 2. Algorithm flow of the proposed method

TABLE I
MAIN PARAMETERS OF THE FMCW RADAR.

Parameter	Value	Unit
Center Frequency	24.125	GHz
Bandwidth	250	MHz
Sampling Frequency	125	kHz
Chirp Duration	1	msn
Unambiguous Range	38.4	m
Distance between antenna	50	cm
Azimuth beamwidth of the antenna	6	degree

IV. METHOD

To clarify the proposed method, the algorithm flow is presented in Figure 2. The output of each radar channel are the digitized IF signal. The first step is the FFT in fast time. The amplitude of the signal is given as input to the target detection algorithm. The detected bins are consolidated and a slow time profile of the target is extracted. The range estimation is achieved as in Equation 7. Then the DPCA algorithm for two channels is applied for time difference and as a result velocity estimation as in Equation 8. The frequency spectrum of the signal is obtained by the FFT in slow time. The Doppler centroid frequency is estimated as it is described in [12]. The estimated range, velocity and Doppler centroid frequency is used to obtain the 2D matched filter and convoluted with the channel output to extract the final ISAR image.

V. RADAR SYSTEM DESIGN

The radar system parameters are shown in Table I. The K band and 250MHz bandwidth is chosen as this is the allowed ISM band for the traffic monitoring radars.

The range resolution is 0.6m and each chirp is digitized up to 128 real samples to fulfill two specifications of the system: The first is the maximum range to be 38.4m. Second is that

the vehicles is associated with the lanes accurately. Typically, the width of the lanes is approximately 3.5m.

The chirp duration shall be chosen so as to avoid aliasing in the Doppler bandwidth. The maximum target velocity is assumed to be 50m/sn that corresponds to 977Hz in K band. So the sampling frequency is selected so as to achieve a chirp duration of 1msn which yields a 1kHz bandwidth.

The isolation of the target is important for parameter estimation and is achieved if there is a range difference between the targets. So the number of targets in the same range must not be more than one. That is why, a narrow beam antenna in the azimuth direction shall be chosen for the proposed system. The maximum value of the illuminated azimuth footprint with the parameters in Table I is approximately 4m that is comparable with the length of a single car and is a distance that is less likely long enough to include multiple vehicles.

Another system parameter is the distance between the radar antenna phase centers. For a maximum velocity of 50m/sn, and velocity accuracy of 10% and 1msn chirp duration the minimum distance between the radars shall be at least 50cm. We emphasize that this can be improved by fine estimation of the frequency.

The multiplexing of two LFM's can be achieved in a time division or a frequency division manner. A frequency division without interference is achievable as the bandwidth of the received signal is quite narrow comparing with the sweep frequency. In this case, the frequency stability of the sensors shall be guaranteed not to interfere to each other. In the time division case, the multiplexing can be done between the chirps and within the chirps. The previous one may result in poor cross correlation of the received signal while the second one requires a highly accurate synchronization. For this purpose frequency division multiplexing with synchronizing the initial timings of the chirps is followed.

VI. RESULTS

In this section, the results of the simulation and on site experiments will be discussed. The parameters of both is given in Table I.

To show the performance of the method, 6 point targets are located to simulate a vehicle of size 3m by 3m. The steps of the simulation is as follows: Raw data simulation, FFT in fast time, velocity estimation, FFT in azimuth, DCF estimation, reference function construction and multiplication with the raw data. The output of the FFT and the final ISAR image are presented in Figure 3. In the upper part of the figure, it is seen that the target is smeared in azimuth as this is an unfocused radar image. However in the below image, the point targets are focused that results in accurately estimating the length of the target.

In the next step, the performance of the proposed algorithm for the real traffic case is assessed by using two commercially available FMCW radars that are located along the road direction. Both of the radars transmit the LFMCW signal with the parameters that are given in Table I. The radars are synchronized to each other by using the same clocks to initialize the waveform. The received data is digitized after low pass filtering.

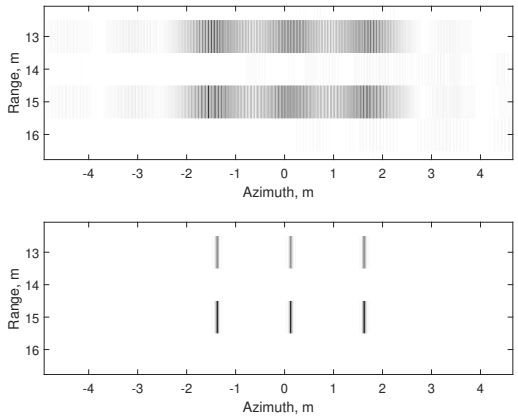


Fig. 3. Simulation results for the 6 point targets case. Up: The fast time FFT of the received signal. Down: After the proposed method is applied

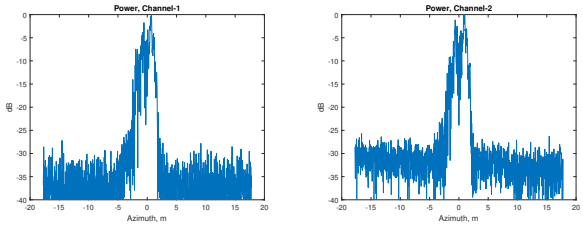


Fig. 4. Azimuth profile of the vehicle for two channels

In Figure 4, the power distribution of a vehicle versus the azimuth is shown for two channels. The estimated velocity for this target is 22.9m/sn. The azimuth time of the target is multiplied by the estimated velocity to extract the azimuth axis in meter. As it is shown, it is not easy to estimate the length of the target due to the range and antenna pattern dependency of the azimuth profile. A target at a different range will have a different radar azimuth image that results in poor classification accuracy. The same is the case for different antenna patterns and squint angles.

In the next step, the DCF is estimated and the phase compensation is done. In Figure 5, the power spectrum versus Doppler frequency of the target is shown. The squint angle is estimated to be 2.85 degree and the DCF is 179.1Hz. In the right figure, it is observed that the compensation achieves a centered spectrum.

Finally, the SAR image is shown in Figure 6 with a color map that black represents the highest power. In the upper image, the output of the FFT in range is shown while in the lower part, the output of the 2D matched filtering is shown. The azimuth profile of the vehicle is observed to be shorter in the unfocused image than the focused one. In addition, as the matched filtering is a low pass filter, the background noise is also suppressed.

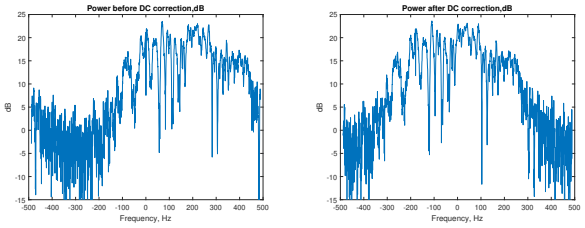


Fig. 5. Power spectrum of the signal. Left: Before DCF compensation, Right: After DCF compensation

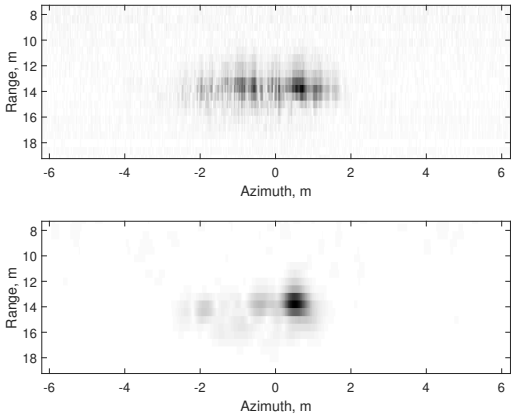


Fig. 6. SAR Image of the vehicle. Up: The fast time FFT of the received signal. Down: After the proposed method is applied

VII. Conclusion

In this work a multistatic FMCW ISAR imaging system is discussed. The proposed system is verified with the simulation and on site experiments. The results show that the proposed radar system is capable of not only range and velocity estimation but also generating the ISAR image. We emphasize that unlike the alternative sensors like cameras, inductive loops and LIDARs, a single side fire radar system is mostly sufficient to monitor a slice of the road without intruding the road. One disadvantage of the system is that a target closer to the radar may prevent to detect a target that is behind. However this performance degradation can be improved by installing the radar to a higher location. It is not easy to quantify the proposed method with the real data. That is why for the future works, a more under control traffic setup that includes moving reflectors can be designed for validation purpose.

References

[1] B. W. Grone, "An Evaluation of Non-Intrusive Traffiffic Detectors at the NTC/NDOR Detector Test Bed", Master Thesis, University of Nebraska, May, 2012
[2] A. Crouzil, L. Khoudour, P. Valiere, D. Nghy Truong Cong, "Automatic vehicle counting system for traffic monitoring", Journal of Electronic Imaging 25(5):051207 · June 2016
[3] P. Asuzu, C. Thompson, "A Signal Model for Simultaneous Range-Bearing Estimation in LFMCW Radar", IEEE Radar Conf. 2019
[4] A. G. Arguello, D. Berges, "Radar Classification for Traffic Intersection Surveillance based on Micro-Doppler Signatures", EuMA 2018
[5] J.M. Muñoz-Ferreras, J. Calvo-Gallego, F. Pérez-Martínez, "Monitoring Road Traffic with a High Resolution LFMCW Radar", IEEE Radar, 2008
[6] S. Sharanov, B. Hill, V. Malik, S. Smith, Robust Dual-Radar-Beam systems and methods for traffic monitorins", Patent, WO/2015/143243, 2015
[7] G. Palubinskas, H. Runge, "Radar Signatures of a Passenger Car", IEEE Geoscience and Remote Sensing Letters, vol. 4, no. 4, October 2007
[8] J. J. M. de Wit, A. Meta, P. Hoogeboom, "FM-CW Synthetic Aperture Radar",IEEE Geoscience and Remote Sensing Letters, vol. 3, no. 1, January 2006
[9] Mehrdad Soumekh, Synthetic Aperture Radar Signal Processing with MATLAB Algorithms, Wiley, 1999
[10] M. Richards, "Principles of Modern Radar", Scitech Publishing, January 31, 2014
[11] O. Dogan, I. Erer, "High resolution radar imaging using GPOF based data extrapolation", IEEE International Geoscience and Remote Sensing Symposium, Alaska Anchorage, 2004
[12] I. G. Cumming, F. H. Wong, "Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation", Artech House, 2005

On Bit Error Rate Analysis In Various Mobility Based Advanced Cellular Communication System

Sutanu Ghosh*, Payel Midya*, Rintu Kumar Gayen†, Kaushik Patra‡, Malay Gangopadhyaya**
Department of Electronics and Communication Engineering
Institute of Engineering and Management
Salt Lake, Kolkata - 700091, India
Email: sutanu.ghosh@iemcal.com*, payel.midya@iemcal.com*, rintu.kumar.gayen@iemcal.com†,
kaushik.patra@iemcal.com‡, malay.ganguly@iemcal.com**

***Abstract*—This paper explores the performance analysis in terms of bit error rate (BER) for different modulations. This estimation is done with the help of various size of control region in advanced cellular communication system. Entire comparative analysis of this paper is elucidated on the basis of different mobility based scenario for downlink (DL) transmission. The simulation is performed using supportive bandwidth of 10 MHz for 2x2 and 4x4 multiple-input-multiple-output (MIMO) configurations. From output results analysis, it is shown that lower mobility environment and lower order modulation has less error to deliver better quality of signal.**

***Index Terms*—Bit error rate, control region, multiple-input-multiple-output, spatial correlation.**

I. Introduction

The recent trends of cellular communication include high data rate applications, reliability, low latency, privacy etc. The advanced generation of cellular communication (like, LTE advanced, 5G, etc) has such kind of capabilities to satisfy the goal of data rate enhancement and improvement of data transmission quality. The progressive path of wireless communication was developed by third generation partnership project (3GPP) [1].

The latest advanced technologies like multiple input multiple output (MIMO) for antenna array configuration, orthogonal frequency division multiplexing (OFDM) for downlink (DL) and single carrier frequency division multiplexing (SCFDM) for uplink (UL) transmission are implemented to obtain the better quality of service. The goal of this paper is to focus the DL transmission technology.

In typical, DL frame structure consists of 10 subframes and each subframe contains 12 or 14 OFDM symbols, which depends upon the applied extended or normal cyclic prefix (CP). Control signals are located in the control region of first c ($c \leq 3$) number of OFDM symbols [2]. These symbols are vigorously varied in each of the subframe. The accessible bandwidth is possible to split into a number of smaller bands in DL using Orthogonal Frequency Division Multiple Access (OFDMA). Those bands are known as sub-channels, and different sets of sub-carriers are assigned to various users to serve the flexibility over the multi-user access technology. These sub-carriers are separately modulated by different modulation techniques like Quadrature phase shift keying (QPSK), 16-Quadrature amplitude modulation (16-QAM) or, 64-QAM modulation [3]. Thus, huge numbers of carriers are added through Inverse Fast Fourier Transform (IFFT) in OFDM technique and that is the main reason behind of the problem of high peak to average power ratio (PAPR). This high PAPR can be reduced using clipping technique. But this clipping technique is responsible for high bit error rate (BER). As quality of the communication system is directly related to error rate, therefore high BER creates massive degradation of system quality. So, the main objectives of this paper is to study the system performance in terms of BER at different modulation techniques.

Another objective of this paper is to investigate the system performance in various mobility based scenarios like low velocity, medium velocity and very high velocity. The recent research on high mobility environment of long term evolution for railway (LTE-R) [4]-[5] is the main motivation of this work to evaluate the error performance on the basis of different modulation schemes [6]. In this work, the size of the control region is considered as 1 and 2.

The main contributions of this work are as follows.

- The comparative study is made in terms of BER for different mobility based scenarios for 10 MHz bandwidth.
- It is also shown the comparison between three different modulation techniques - QPSK, 16- QAM or, 64- QAM.

The remaining sections of this paper are arranged as follows. MIMO based correlation in LTE technology is discussed briefly in Section II. Section III introduces the system model and Section IV shows the performance evaluation. Finally, the paper is concluded in Section V.

II. Overview of spatial correlation of MIMO

MIMO technology is used in cellular communication to serve higher flexibility to adapt more antennas at base station. These multiple antenna pattern forms different antenna array configuration to accomplish efficient data transmission. Conventional MIMO antenna array configurations can be described as 2x2 (it means, both the transmitting and receiving terminals contains two antennas), 4x4 etc. Overall system throughput based on MIMO technology is not only limited upon the number of channels, but also the correlation between those channels is equally important. Feasible range of Correlation

978-1-7281-9615-2/20/$31.00

factor 0 to 1 is used to determine the amount of correlation. Same channel characterization is described through high level correlation with maximum value of 1. On the other hand, self determining characteristics are expressed through low level correlation with minimum value of 0. Kronecker model is very well known model to evaluate MIMO channel performance [7].

Mathematically, Kronecker product [8] of spatial correlation can be expressed between user equipment (UE) and evolved node-B (eNB) as follows.

$$R_{MIMO}^{KRO} = R_i \otimes R_j \tag{1}$$

where, $i, j \in$ UE, eNB; $i \neq j$. The mathematical meaning of the symbol $\otimes$ is Kronecker product.

III. System Description

This section is introduced to describe system model. In cellular architecture, concurrent data transmission are implemented through OFDM technology and MIMO is used to enhance the throughput with robustness [9]. This OFDMA based system architecture contains M number of subcarriers and N number of users where each of them takes L non-overlapping subcarriers. M_g samples are added here to extend OFDM symbols cyclically, so total length can be calculated as M_t= M+M_g samples. Now, u^{th} OFDM symbol can be expressed in the following way:

$$x_u = AD^H m_u \tag{2}$$

where, m_u is a vector includes constellation symbols with dimension of M x 1 to transmit u^{th} symbol sub-carriers. M $\times$ M DFT matrix is expressed by D, A indicates M_t $\times$ M matrix to present the cyclically extend of OFDM symbol. Transmitted OFDM symbol x_u can be expressed as a vector M_t x 1. B is another matrix M $\times$ M_t is applied to remove cyclic prefix. Now x_u with large cyclic prefix over time varying channel is received as:

$$r_u = DB(H_u(t)AD^H m_u + n_u) = H_u m_u + w_u \tag{3}$$

where, $H_u(t)$ indicates the channel impulse response matrix M_t $\times$ M_t, n_u is used to define an uncorrelated complex-valued white Gaussian noise matrix M_t $\times$ 1 with power of σ_n^2 and w_u is used to define uncorrelated complex-valued white Gaussian noise matrix M $\times$ 1 with variance of σ_n^2. H_u can be represented by a diagonal matrix iff the channel is time invariant. On the other hand, non-zero entries will come out from the main diagonal if the channel is time selective and due to this effect inter-channel interference (ICI) will occur in received signal. This ICI can be related to normalized channel Doppler spread, which may be defined as D_s = $f_{dp} T_{sm} M$, where f_{dp} is Doppler frequency and T_{sm} sampling interval. It is possible to control the factor $T_{sm}M$ using interpolation component of i. This interpolation effect with Doppler spread at given velocity of mobile receiver v_m can be written as :

$$D_s^i = f_{dp}(T_{sm}M)^i = f_{dp} i (T_{sm}M) = (f_{ca}^i T_{sm} M v_m)/c \tag{4}$$

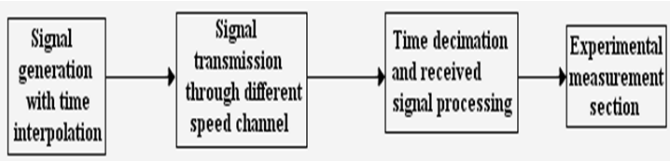


Fig. 1. System model

TABLE I
List of necessary parameters

Parameters	Value
Subframe Size	10 (This 10 Subframes form a Radio Frame with the duration of 10 ms)
Maximum Doppler Shift	12.133, 121.133 and 275 Hz
Utilized Bandwidth for this work	10 MHz

where, carrier frequency and velocity of light are f_{ca} and c, respectively. In this work, i is taken as more than 1. The system model of system architecture is given in Fig. 1.

IV. Simulation Results

Performance evaluation is shown using MATLAB and the output results are validated through Qualnet software [10]. The results is based on different velocity of moving objects as shown in different graphical result analysis. LTE standard frequency of operation is taken as 2.6 GHz. This is assumed that the system is perfectly time and frequency synchronized. Signals of existing system model are propagated through different discrete paths, which have average gain vector ranging from 0 to -6 dB. LTE downlink channel information can be described through Physical Control Format Indicator Channel (PCFICH) to parameterize the size of control region. First 3 OFDM symbol (sequentially 1, 2 and 3) in each subframe is used to make an availability of downlink control signal. This part of subframe is described as control region in the system model of this paper. Different system parameters used for simulation are listed in Table I.

For 64 QAM, 16-QAM modulation in Fig. 3 and Fig. 4, it is seen that the minimum BER is found at 20 dB SNR for lowest mobility UE (1.4 m/s) with 2x2 antenna array. For QPSK modulation - Lower BER (0.00000741) is observed in Fig. 2 at 10 MHz bandwidth for lowest mobility UE (1.4 m/s) with 4x4 antenna array. This lower level of BER is saturated from 16 dB to 25.11 dB and finally it is dropped to 0 value beyond that range. On the basis of this observation, QPSK modulation is found to be lower optimum level of BER value at 20 dB SNR for 10 MHz bandwidth and 2x2 antenna array. The performance is improved using more number of antennas at the both end of cellular system.

For 64 QAM and 16 QAM modulation in Fig. 6 and 7, it is easily found that the minimum BER is observed at 20 dB SNR for lowest mobility UE (1.4 m/s) with 2 $\times$ 2 antenna array.

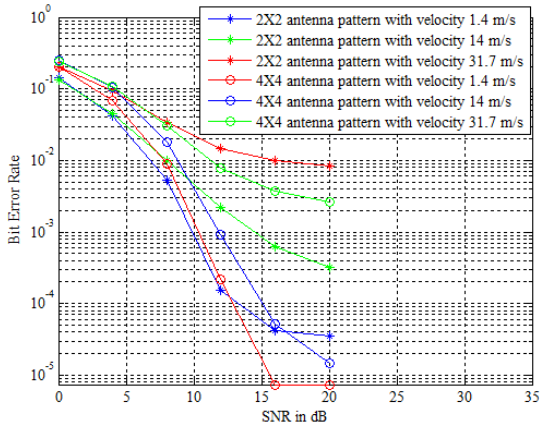


Fig. 2. BER plot for 10 MHz Bandwidth and 1 control region for QPSK modulation technique

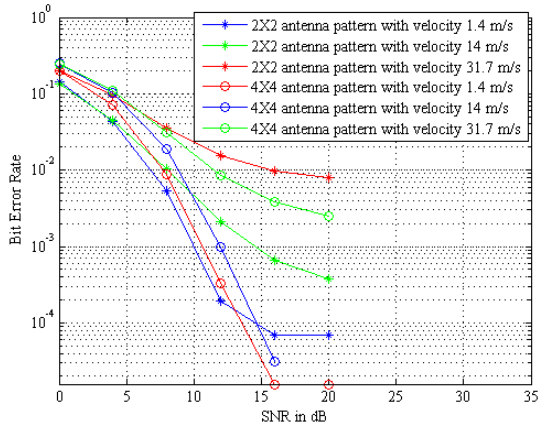


Fig. 5. BER plot for 10 MHz Bandwidth and 2 control region for QPSK modulation technique

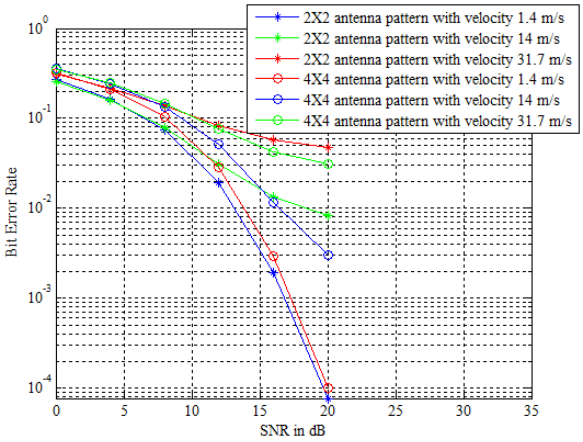


Fig. 3. BER plot for 10 MHz Bandwidth and 1 control region for 16-QAM modulation technique

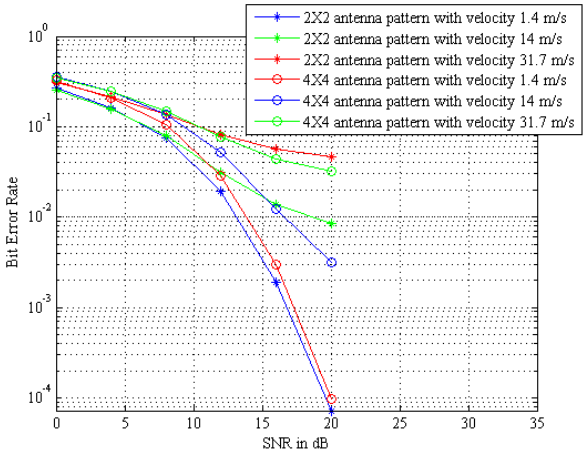


Fig. 6. BER plot for 10 MHz Bandwidth and 2 control region for 16-QAM modulation technique

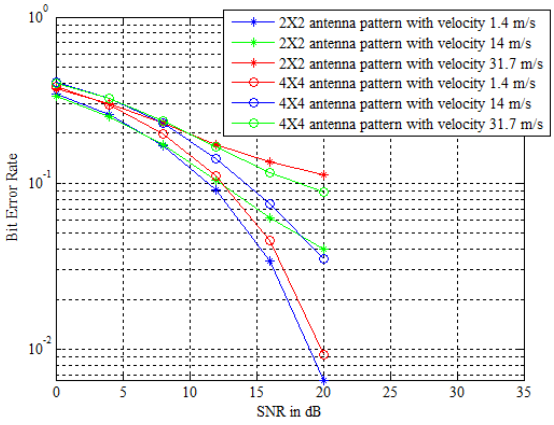


Fig. 4. BER plot for 10 MHz Bandwidth and 1 control region for 64-QAM modulation technique

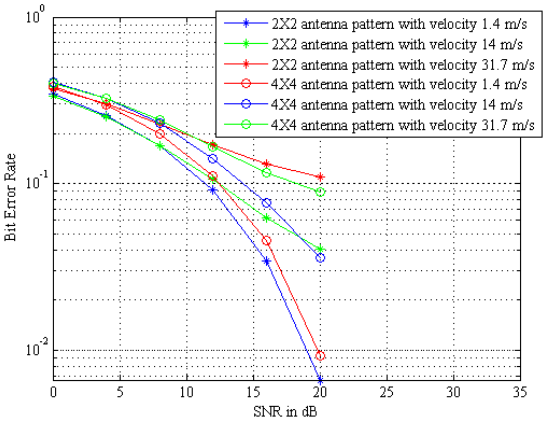


Fig. 7. BER plot for 10 MHz Bandwidth and 2 control region for 64-QAM modulation technique

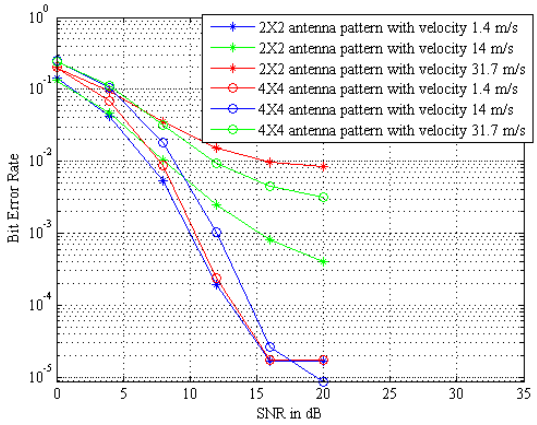


Fig. 8. BER plot for 10 MHz Bandwidth and 3 control region for QPSK modulation technique

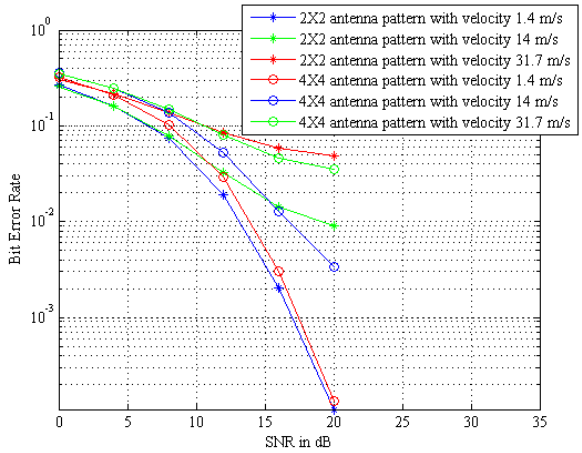


Fig. 9. BER plot for 10 MHz Bandwidth and 3 control region for 16-QAM modulation technique

Due to better channel estimation in QPSK modulation (in Fig. 5), almost 0 level of error is observed in this modulation for medium velocity equipment (14 m/s) at 20 dB SNR with 4x4 antenna array. On the basis of this case study at 20 dB SNR, lower optimum level of BER value for QPSK modulation is observed for 4×4 antenna array.

For 64 QAM modulation and 16 QAM in fig. 9 and 10, it is studied that the minimum BER is found at 20 dB SNR for lowest mobility UE (1.4 m/s) with 2×2 antenna array.

The BER performance of QPSK (in Fig. 8) is almost same like control region 2.

V. Conclusion

In this paper, the BER performance is shown with respect to different modulation techniques and different antenna configurations in 3 different mobility scenarios. On the basis of data analysis, it is observed that for increasing number of antenna array, lower order modulation has negligible amount of error. On the other hand, higher order modulation can perform better for 2×2 antenna array. Higher velocity environment is performing with higher error than low and medium velocity environment.

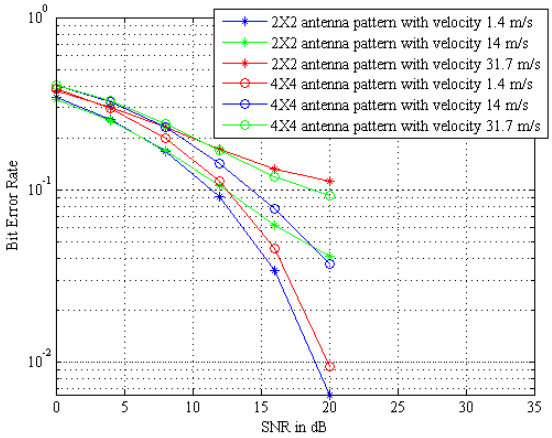


Fig. 10. BER plot for 10 MHz Bandwidth and 3 control region for 64-QAM modulation technique

References

[1] "3GPP - The mobile broadband standard," *available at : http://www.3gpp.org*.

[2] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe and T. Thomas, "LTE-advanced: next-generation wireless broadband technology," *IEEE Wireless Communications Magazine*, vol. 17, no. 3, pp. 10-22, Jun. 2010.

[3] D. Martín-Sacristán, J. F. Monserrat, J. Cabrejas-Penuelas, D. Calabuig, S. Garrigas, and N. Cardona , "On the way towards fourth-generation mobile: 3GPP LTE and LTE-advanced," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, DOI: 10.1155/2009/354089, pp. 1-10, Sep. 2009.

[4] K. Guan, Z. Zhong and B. Ai, "Assessment of LTE-R Using High Speed Railway Channel Model," *Third International Conference on Communications and Mobile Computing*, pp. 461-464, Qingdao, China, Apr. 2011.

[5] W. Luo, R. Zhang, and X. Fang, "A CoMP soft handover scheme for LTE systems in high speed railway," *EURASIP Journal on wireless Communications and Networking*, vol. 2012, DOI : 10.1186/1687-1499-2012-196, pp. 1-9, Jun. 2012.

[6] S. Ghosh, "Performance Evaluation of Different Coding and Modulation Scheme in LTE Using Different Bandwidth and Correlation Levels," *Springer Wireless Personal Communications*, vol. 86, no. 2, pp. 563-578, Jan. 2016.

[7] K. Werner, M. Jansson, and P. Stoica, "On estimation of covariance matrices with Kronecker product structure," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 478-491, Feb. 2008.

[8] D. Molteni, M. Nicoli, and U. Spagnolini, "Performance of MIMO-OFDMA systems in correlated fading channels and non-stationary interference, " *IEEE Transactions on Wireless Communications*, vol. 10, no. 5, pp. 1480-1494, May. 2011.

[9] T. T. Nguyen, V. T. Nguyen, T. A. Hoang, T. M. Nguyen, and K. L. Dang , "Performance comparisons of detector algorithms for high data rate MIMO OFDM systems in frequency selective fading channel," *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 1-5, Ho Chi Minh City, Vietnam Dec. 2012.

[10] Qulanet Simulator, Scalable Network Technologies, available at: www.qualnet.com, 2012.

Web Accessibility Analysis on Government of India Websites based on WCAG

Abhirup Sinha
Analyst, Tata Consultancy Services Limited
sinha.abhirup@gmail.com

***Abstract*— Information technology has become a part of day to day life. In every part of day to day information technology has its intervention. Government activities and services are also part of this digital revolution. Many of the Governments around the world had already provided the services in digital form and some are in the process of change. That means the services offered by Government are available through web-applications and websites and users do not go to physical offices to avail these services. These users include those as well who are differently abled and need specific help to perform tasks. These disabled users are mostly ignored while developing digital applications. So, there are lot of websites where these specially challenged people are not able to do usual work as per expectation. This study evaluates Web Accessibility compliance and provides over all report on how Government of India's websites are developed to handle World Wide Web Consortium's Web Accessibility standard. Here reports have been generated based on tools and Web Content Accessibility Guideline (WCAG) 2.0 are followed. After analyzing accessibility scores and issues, the result says, websites of government of India still need lot of improvement to reach standard web accessibility standard.**

Keywords—Web Accessibility, WCAG 2.0, WCAG 2.1

I. INTRODUCTION

Day by day WWW is getting importance in daily life and gradually it became part of our daily life. Starting from private, public and even government sector also implementing all their customer faced service in digital platform. Government of India has taken Digital India initiative to give seamless digital service to all citizen across all jurisdictions to make life easier [1]. While making everything digital and all web based service, keep in mind about web accessibility compliance and need to make sure to give equal right and platform to disabled users to access all websites.

At present around 15% of world population is living with disabilities and this percentage results around 1 billion [2][3]. In case of India, current population is around 1.3 Billion. As per last census in 2011, India was having 2.21% of total population, who was having some sort of disabilities [4]. With a simple calculation based on current population, around 26 million people are living with some sort of disabilities in India. Considering this huge number of disabled people in India, its easily understandable how important is to give equal privilege to these users on daily web life. Government websites are having very crucial role here on web platform to give equal experience to disabled users.

This paper is structured as follow: second section (II. Literature Study) states about existing works on related area, third section (III. Web Accessibility Standard and Guidelines) is about background of web accessibility and guidelines we have, fourth section (IV. Research Methodology and Tools used) describes research methodology and tools used for this study conducted in this paper, fifth section (V. Result, Analysis and Discussion) is all about analyzing result, and in sixth section (VI. Conclusion) paper has been concluded followed by Future work.

II. LITERATURE STUDY

As Web Accessibility is very important section in Internet era so there are good numbers of research work or resources are available. All these case studies are based on either automated tool, manual evaluation process or by both. As per Mark Rogers [6] most of the countries follow straight away WCAG 2.0 as standard in Web Accessibility, couple of others use their own guideline which is prepared based on WCAG 2.0. As per one of these studies done by Norman [8], there are lot of issues in USA state government websites also. There evaluation has been done based on WAVE automated tool and only errors have been counted for analysis and not alert. In a study conducted by Beaudin [9], total 66 websites were evaluated form 3 countries (22 from each). The result of the study states that 86% Government

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

websites of USA, 23% government website of UK and 54% Australian Government websites are satisfying WCAG web accessibility compliance. A study by Tan, C.W. [10] gives us idea how good website can build trust of citizen towards government. An article by Faries, A., Khalifa, H.S., Razgan, M.S., Al-Duwais, M. [11] shows accessibility compliance coverage on top Saudi e-government services with suggested framework that can be implemented not only to enhance accessibility coverage but also to make better usability. Detailed case study by Isa, W.A.R.W.M., Suhami, M.R., Safie, N.I., Semsudin, S.S.[12] on 155 Malaysia Federal websites concluded with result where lots of Accessibility issues were identified which causes violation of WCAG 2.0. A study done by Mourad, B., Kamoun, F. [13] on Dubai e-governance websites states most of the websites are still not fully accessible and there are lot of scope for improvement. Also, this research result states that most of the accessibility issues are related to non-text equivalent content. Similar type of case study also conducted by Adepoju, S.A., Shehu, I.S., Bake, P. [14], Karaim, N.A., Inal, Y. [15], Ismailova, R. [16], Akgul, Y., Vatansever, K. [17] where suggestion has been given to improve websites based on web accessibility standard. Joanne Kuzma did a detailed study [7] on multiple countries e-government websites and concluded with result where it's clear that most of the accessibility errors are related to alternative non-text content. A study by Silvia [18] states that learning websites of Mathematics, Geometry and Physics are having accessibility issues those are mainly navigation, color contrast and alternative text related.

III. WEBACCESSIBILITY STANDARD AND GUIDELINES

The World Wide Web Consortium (W3C) develops standard for website development. While defining the standard, they also defined how to follow web accessibility standard in websites. They have taken Web Accessibility Initiative (WAI) to standardize process and implementing guidelines that to follow while developing any website.

Web Content Accessibility Guidelines (WCAG) is developed by organizations as well individuals who are related to W3C. The goal of this guideline is to provide standard 'do and don't' for web content development to the needs of individuals, organizations, and governments internationally [5]. There are multiple versions of WCAG guidelines. WCAG 1.0 was published by W3C and became a recommendation on 5th May 1999. WCAG 1.0 later followed by WCAG 2.0, which got published on December 2008 and still community is continuing to refer that. In June 2018, WCAG 2.1 got published and W3C recommended to follow that. There is not much difference between WCAG 2.0 and 2.1. WCAG 2.1 contains all 2.0 guidelines and in addition to that there are few guidelines which supports mobile web accessibility.

WCAG guidelines have been categorized in three parts or conformance. These are in Table 1:

Table 1: WCAG Conformance level

Conformance level	Description
A	These are most basic web accessibility features, and these should be followed for any website development.
AA	This deals with major accessibility barrier and developer should follow standard coding procedure to satisfy compliance with AA.
AAA	This conformance is like good to have. These are not mandatory but if webpage could be compliance with AAA then will be more accessible.

As per industry standard, best practice is to meet the target to comply with all A, AA and few AAA (if possible) guidelines. This case study will mainly put focus on A and AA conformance level guidelines.

IV. RESEARCH METHODOLOGY AND TOOLS USED

Intension of this case study is to do an analysis on Indian Government websites and to understand that how much they are compliant to W3C Web Accessibility guidelines. Here automated web accessibility tools are used, mentioned in Table 2, for validation and generating reports.

Table 2: Evaluation tools

Tools	Description
Lighthouse	This is an open source tool developed by Google which can evaluate web page form its Performance, Accessibility perspective.
Wave	This is also a free tool which can be added as add-in to our browser and it can evaluate and generate report from Web Accessibility standpoint.

After doing analysis result is published and that provides generic recommendations also. By Following those recommendations, websites could be more accessible for disabled users.

Lighthouse is an open source tools that developed by Google and it is available on developer tools of Google chrome browser. This tool will give over all percentage on web accessibility. Also, it will show list of things got passed and failed. This tool is having a link also where developer can refer how to make website more accessible. Second tool, WAVE is a add on tool and can be added with Google Chrome or Fire Fox browser. By using this tool, you can get more detailed report and pictorial view of erroneous area or component.

There are multiple other tools also in market (both open source and licensed) using which automatic evaluation can be done on any web site from Accessibility standpoint. But need to keep in mind one thing – not a single tool is 100% accurate and manual validation should be done on each component on web page to understand or find all accessibility related issues. However, automation tools are very helpful here to understand overall picture of web accessibility compliance on any website.

V. Result, Analysis and Discussion

There are lot of free Web Accessibility Evaluation tool available in market. However, it's always advisable to use a tool that is from authentic source. In this Result analysis Google's Lighthouse and AIM's WAVE will be used considering usability, reporting structure and authentic source of these tools.

A. Google Lighthouse Accessibility evaluator

As it was mentioned earlier that Google Lighthouse is a free tool that by default available on Google Chrome browser's developer window. Using this tool, I have evaluated 5 random page of websites including Home page. As home page is the primary page of any website so have included that on all my analysis. I have taken 10 Government of India websites. Light house evaluator is used on all websites and noted Accessibility score based on the result got from tool. Referring below Table 3 could see the scores for each page of each websites. As per this Lighthouse tool if any page scores 100 then that is fully accessible. Also, what can be said is, any score more than 90 is good and indicates that if couple of issues got fixed then that page will be good from Accessibility standpoint. A score below 90 and above 80 represents partially accessible and anything below 80 is not accessible. In below Table 3, Page 1 indicates Home page. In present time, #makeinindia is flagship program of Government of India. However, if you see the Page 1 Web Accessibility (Ax) score of makeinindia.com, its 59, which is very poor and is indicating, this website is not at all accessible. Same goes to another popular and mass visited website passportindia.gov.in, which scored 66, better than makeinindia.com but not a promising score. These two websites scored below 70 and what could be said is coding issues must be fixed to make these websites accessible. Detailed issue discussion will be there on Wave tool report section. There are five homepages where Lighthouse score is between 80-89. These are partially accessible, and after fixing issues, it will be well enough from accessibility perspective. There are three homepages where having 90+ score and those are, indiapost.gov.in, epfindia.gov.in and upsc.gov.in. These home pages are almost accessible and need minimal fixes to make those fully accessible. So, considering homepage results of ten websites what could be said is, 20% home page is not accessible, 50% is partially accessible and 30% is almost accessible. Now on the Figure 1, where average Ax score is there considering all five pages of listed websites. As per Figure 1, only two website's average score is 90+ and those are indiapost.gov.in and upsc.gov.in. there are six websites which scored between 80-89 range and there are two websites where average score is below 70. These scores could be changed if evaluation would be done on more pages for each website.

Also, it's always advisable to do a manual evaluation to find more accessibility issues. In manual validation, all screen reader related issues can be identified, which are not possible to get in any automated tools.

Table 3: Lighthouse Ax Score

Lighthouse Report	Web Accessibility Score				
Websites	Page 1	Page 2	Page 3	Page 4	Page 5
incometaxindiaefiling.gov.in	84	74	81	90	88
indiapost.gov.in	91	92	91	91	89
passportindia.gov.in	66	91	78	81	81
epfindia.gov.in	91	72	39	77	50
ssc.nic.in	81	82	82	83	82
incometaxindia.gov.in	82	90	88	91	88
upsc.gov.in	94	92	92	91	87
india.gov.in	83	83	83	87	83
mygov.in	83	87	87	84	84
makeinindia.com	59	65	65	52	65

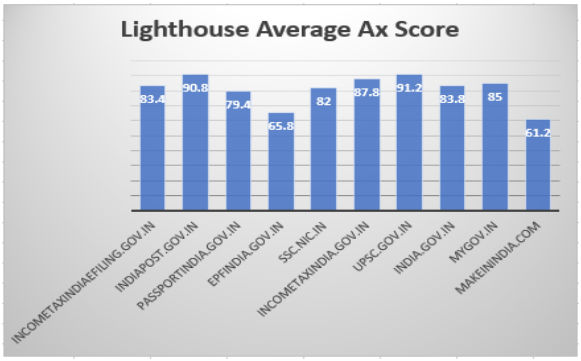


Figure 1: Lighthouse average Ax score

B. WAVE Web Accessibility Evaluator

Using WAVE tool, same list of websites has been evaluated. Here analysis has been done only on home page of those websites. WAVE is a tool which will provide a graphical representation of location of issues. It will indicate all components where the web accessibility violation is. Also, it provides a dashboard where could be seen overall issues. Referring Figure 2, the dashboard can be seen where over all evaluation has been represented. Clicking on 'Details' tab on Wave dashboard, could get details of issues or error. Sample detailed view can be seen on Figure 3.

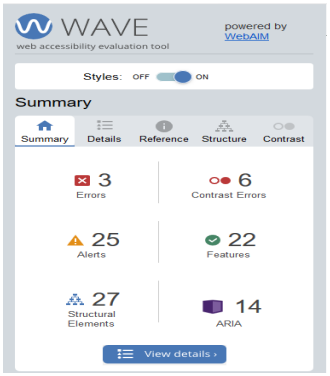


Figure 2: WAVE Dashboard

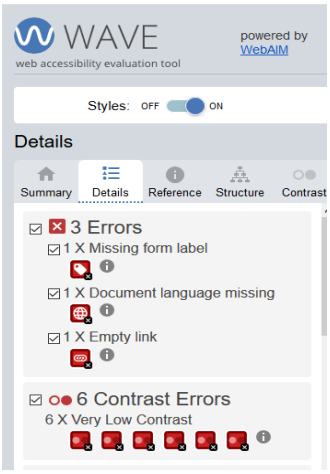


Figure 3: WAVE Error details

While evaluating website's homepage it was observed that all of those are violating certain success criteria. All these webpages are having errors mainly with Color contrast, Missing Label, Page language, Empty Link, Alt Text and Empty Button. In Table 4, can be seen which website is having how many issues related to these errors. Referring Table 4, it can be said that all websites have one common error and that is Color Contrast issue, which is very much critical in Web Accessibility. Figure 4 indicates how many total Accessibility errors are there in listed webpages.

Table 4: WAVE Report

Websites	Contrast	Missing Label	Page Language Missing	Empty Link	Alt text	Empty button
incometaxindiaefiling.gov.in	6	1	1	1	0	0
indiapost.gov.in	8	0	0	0	0	0
passportindia.gov.in	15	3	0	16	20	1
epfindia.gov.in	70	0	1	1	0	0
ssc.nic.in	2	10	0	0	0	3
incometaxindia.gov.in	8	0	0	2	3	3
upsc.gov.in	17	0	0	3	0	0
india.gov.in	32	0	0	5	8	0
mygov.in	80	3	1	25	32	0
makeinindia.com	5	0	0	218	1	0

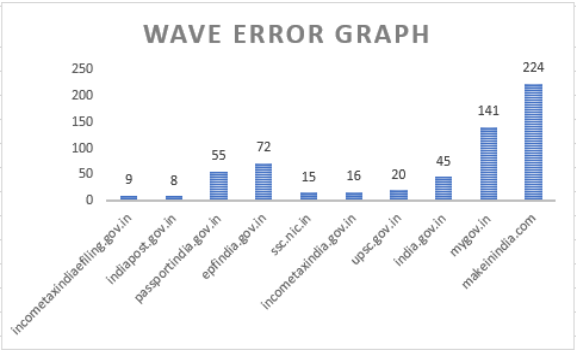


Figure 4: Webpage Error count

After summarizing the evaluation done by WAVE tools, will get Figure 5, that states Issue Occurrence count with Errors. It clearly states that, in all ten listed websites, most of the issues are with Color Contrast and Empty Link. There are total 243 Contrast violation and 271 Empty Link issues. Apart from these two, 64 issues are due to ALT Text, which results bad web accessibility.

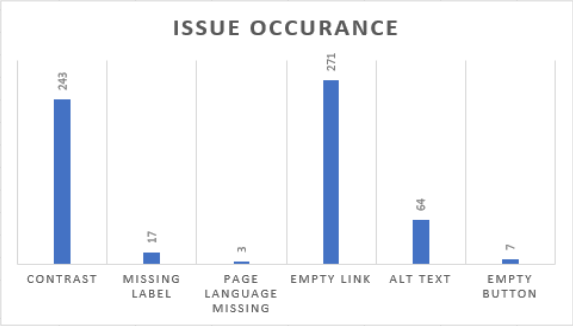


Figure 5: Wave Issue Occurrence

Table 5: WCAG Violation

Issues	WCAG Violation
Contrast	1.4.3 Contrast (Minimum) (Level AA)
Missing Label	1.1.1 Non-text Content (Level A) 1.3.1 Info and Relationships (Level A) 2.4.6 Headings and Labels (Level AA) 3.3.2 Labels or Instructions (Level A)
Page Language Missing	3.1.1 Language of Page (Level A)
Empty Link	2.4.4 Link Purpose (In Context) (Level A)
Alt text	1.1.1 Non-text Content (Level A)
Empty button	1.1.1 Non-text Content (Level A) 2.4.4 Link Purpose (In Context) (Level A)

Table 5 states that what are the WCAG guidelines violated here. There might be a chance of violating same guidelines for different issues. Errors related to Missing Label, Alt text and Empty button – they all are under one common WCAG violation, which is 1.1.1 Non-text Content (Level A). Also, Table 5 states that there are two AA conformance level violation - 1.4.3 Contrast (Minimum) and 2.4.6 Headings and Labels; along with five Level A violation. As it was already mentioned in Table 1, all these A and AA violation should be fixed to meet web accessibility standard. Apart from all these, it was also observed that there are Marquee on couple of websites, which is also very bad practice and it violates 'WCAG 2.2.2 Pause, Stop, Hide'.

Now will discuss briefly about top three errors were observed on pages, what these errors are and what to do to meet Web Accessibility standard.

- Contrast or Color Contrast
  - It means: Contrast of color between foreground and background is very low.
  - It Matters because: Proper Color contrast is very essential for Low vision user group.
  - Can be fixed by: Increasing color contrast between foreground and background. As per standard, color contrast ratio should be 4.5:1 for small text and 3:1 for large text and images.
- Empty Link
  - It means: Link is not associated with any text.
  - It matters because: User will not understand the purpose of the link.
  - Can be fixed by: Providing descriptive text on that link so that user can understand the purpose of the link or destination of the link if they are putting focus on it while using screen reader.
- Alt Text
  - It means: Alternative text is missing for decorative or non-decorative image
  - It matters because: User will not understand the content of any image with screen reader if there is no alt attribute with that image. Also, if the image is not available due to some reason, then also user will not understand the purpose or meaning of that image.
  - Can be fixed by: Providing proper descriptive and crisp alt attribute for all non-decorative image and at least blank alt attribute for decorative image

As WAVE doesn't give any accessibility score. So, to compare both WAVE and Lighthouse result, need to tally Lighthouse score with 'inverse' Error count of WAVE. In below Table 6, listed top 5 website in descending order with respect to Lighthouse score and ascending order with respect to WAVE error count. Also, websites are marked with green color, those are common in both Lighthouse and WAVE report. There are three websites common for both WAVE and Lighthouse. As these all automated tools evaluate code in different procedure

and manner so it's quite natural that end result will be having little bit of discrepancies. However, here one significant point is, in my study as 3 out of 5 websites are common in both the report so it can be said that there is not much discrepancy in Lighthouse and WAVE report.

Table 6: Lighthouse and WAVE result comparison

Lighthouse	WAVE
upsc.gov.in	indiapost.gov.in
indiapost.gov.in	incometaxindiaefiling.gov.in
incometaxindia.gov.in	ssc.nic.in
mygov.in	incometaxindia.gov.in
india.gov.in	upsc.gov.in

This case study shows that there is scope of improvement in all these Government of India websites. As in India almost 26 Million citizens are living with at least some sort of disability, so serious and sincere steps should be taken towards web accessibility. In a discussion about mass and large-scale digitalization from Government standpoint, expectation is people will come together to be 'digital' rather than standing on human queue. And in this manner web community should take care of our disable citizens too. Neglecting privilege of disabled people will not help us at all to reach our goal and to be fully digital. From human rights standpoint also its not legal to stop specially challenged users to use websites just because those websites are lack of accessibility compliant. In all first world country like USA, UK, EU, Canada they are having very strict policy and all public and Government agencies follow those. They also have grievance cell for non-accessibility compliance and penalty procedure. In India also WCAG guidelines should be followed strictly to make any websites well accessible. It's not a technology to implement, but a regular practice which can make life easier for some special people.

There should be a proper plan and a road map to implement and make website accessible. That approach should have three parts, Grass-root level development, Continuous monitoring and Grievance & Feedback. In grassroot level, WCAG Accessibility Guidelines could be included in curriculum in school or university level to make future Developer aware of it. Students can also be trained on how to evaluate any website from Accessibility standpoint. More next generation will be aware of web accessibility, slowly and gradually problem will be resolved. In Continuous monitoring, there should be a trained team who will be checking websites and will send note on area of improvement to development company. As in current state, India is having booming Information Technology sector so this will be very easy to implement. With proper plan it's not impossible to make all major websites to be accessible. Third one would be Grievance & Feedback where; a dedicated cell will be there where user can log complain or provide feedback on how website is doing from accessibility point of view. After that all those feedbacks and complain will be informed to Developer of that website through 'monitoring' team. Also 'penalty on violation' can be implemented in future so that all organizations also keep importance of web accessibility in their mind. There are lots of free evaluation tools which gives not only report but suggestion on how to fix issues. Additionally, main thing is WCAG guideline is free to use or implement. So, there should be nothing stopping us to make any website accessible.

VI. CONCLUSION

This detailed case study and evaluation result shows that India's Federal Government websites are partially accessible, and not much effort is needed to make those websites fully accessible. Government Organization might need to rethink on design in some cases where user interface and page design needs to be reviewed. However, those are really not an impossible thing where Information Technology professional are so strong in India. Government of India is already having a special campaign named "Sugamya Bharat Abhiyan" for differently abled people. Web accessibility also must be taken care as a part of this program and need heartful presence of all people of India to implement this. Our small but steady initiative can make special people's web life better and hassle free.

VII. FUTURE WORK

In future there is a plan to perform detailed study on a larger number of Government of India websites. Also, statistical analysis will be performed on result along with automatic and manual evaluation. .

REFERENCES

[1] https://www.digitalindia.gov.in/content/vision-and-vision-areas

[2] https://www.un.org/development/desa/disabilities/resources/factsheet-on-persons-with-disabilities.html

[3] https://www.who.int/disabilities/world_report/2011/report/en/

[4] https://enabled.in/wp/disabled-population-in-india-as-per-census-2011-2016-updated/

[5] https://www.w3.org/WAI/standards-guidelines/wcag/

[6] Mark Rogers (13 November 2012). "Government Accessibility Standards and WCAG 2.0". Powermapper.com. Retrieved 15 December 2014

[7] Joanne Kuzma University of Worcester, Global E-government Web Accessibility: A Case Study https://core.ac.uk/reader/49618

[8] Norman E. Youngblood, Susan A. Youngblood, User Experience and Accessibility: An Analysis of County Web Portals, pp. 25-41, 2012

[9] Beaudin, D., A Content Analysis Of Disability Access On Government Websites In Australia, The United Kingdom And The United States, A Master's thesis, the faculty of the School of Information and Library Science of the University of North Carolina at Chapel Hill, North Carolina November, 2001.

[10] Tan, C.W., Benbasat, I., Cenfetelli, R.T.: Building citizen trust towards e-government services: do high quality websites matter?. In: Proceedings of the 41st Annual on Hawaii International Conference on System Sciences, pp. 217–217. IEEE (2008)

[11] Al-Faries, A., Al-Khalifa, H.S., Al-Razgan, M.S., Al-Duwais, M.: Evaluating the accessibility and usability of top Saudi e-government services. In: Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance, pp. 60–63. ACM (2013)

[12] Isa, W.A.R.W.M., Suhami, M.R., Safie, N.I., Semsudin, S.S.: Assessing the usability and accessibility of Malaysia e-government website. Am. J. Econ. Bus. Adm. 3(1), 40–46 (2011)

[13] Al Mourad, B., Kamoun, F.: Accessibility evaluation of Dubai e-government websites: findings and implications. J. E-Gov. Stud. Best Pract. 2013, 1–15 (2013)

[14] Adepoju, S.A., Shehu, I.S., Bake, P.: Accessibility evaluation and performance analysis of e-government websites in Nigeria. J. Adv. Inf. Technol. 7(1), 49–53 (2016)

[15] Karaim, N.A., Inal, Y.: Usability and accessibility evaluation of Libyan government websites. Univ. Access Inf. Soc. 18(1), 207–216 (2019)

[16] Ismailova, R.: Web site accessibility, usability and security: a survey of government web sites in Kyrgyz Republic. Univ. Access Inf. Soc. 16(1), 257–264 (2017)

[17] Akgul, Y., Vatansever, K.: Web accessibility evaluation of government websites for people with disabilities in Turkey. J. Adv. Manag. Sci. 4(3), 201–210 (2016)

[18] Silvia B, Laura M, David V, Gloria V; Accessibility Evaluation of Web Content that support the mathematics, Geometry and Physics's teaching and learning; IEEE ICALT, 2019

Congestion-aware Data Acquisition with Q-learning for Wireless Sensor Networks

Praveen Kumar Donta, Tarachand Amgoth and Chandra Sekhara Rao Annavarapu
Department of Computer Science & Engineering
Indian Institute of Technology (Indian School of Mines), Dhanbad Jharkhand, India
praveeniitism@gmail.com, {tarachand, acsrao}@iitism.ac.in

***Abstract*—In wireless sensor networks (WSNs), congestion is one of the challenging issues. It degrades the network's performance in various operating variables such as throughput, latency, energy consumption, packet loss, lifetime of the network and etc. Congestion occurs in WSN when the rate of sensor data flow outreach the channel or buffer capacity. The congestion is controlled in two ways in WSN, such as controlling network traffic or efficiently managing the resources. This paper performs the resource control mechanism by providing the alternative path towards to base station using a Q-learning for congestion alleviation. This congestion-aware data acquisition (CADA) mechanism initially identifies the congestion node (CN) where the nodes buffer occupancy ratio is higher. Further, we recognize the proper next node to construct the dynamic alternative route to the base station. The CADA is evaluated in various network conditions by comparing it with recent congestion-aware algorithms. The simulation tests show that the CADA efficiently ameliorates the congestion and enhances the performance across multiple performance metrics.**

***Index Terms*—Congestion Control, Wireless Sensor Networks, Data Collection, Q-learning, Resource Control**

I. Introduction

Wireless sensor networks (WSNs) are the most auspicious technology due to its appealing features such as scalability, cost-effective, low-complexity, self-organize, ease of use, and deployment. WSNs are composing with an immense set of tiny sensor nodes that furnished with a communication component, limited storage capacity, sensing, and processing unit [1]. Sensor nodes (SNs) observe the phenomenon by monitoring the environment continuously and transmit it to the sink. The limited communication bandwidth of the sensor nodes sends the data either direct (sink node/base station (BS) within range) or multi-hop manner. Some of the WSNs applications such as smart city applications, battlefield monitoring, target tracking, home automation, and environmental monitoring, etc., are generated a massive amount of continuous data, which increases the network traffic [2]–[4]. Because of a large number of nodes and multi-hop transmissions, when the sensor nodes' traffic exceeds the available capacities of the node buffer or channel contention, it is usually uncontrollable, and it creates the congestion in WSNs [5], [6]. It happens anywhere in the WSN such as source node, near to sink node, or the region where more number of SNs are deployed.

In WSNs, congestion is an increasingly significant challenging issue, because it has more impact on various QoS parameters. It mainly degrade the energy wastage, throughput, packet loss, packet delivery rate (PDR), and latency [7]–[9]. Congestion control (CC) is one of the most striking and challenging issue in WSNs because it is quite disparate than other wireless networks [10]. There are several CC approaches for WSNs are discussed in the literature [11]. The CC mechanism in the literature is categorized into three types i.e., resource control, traffic control, and hybrid approaches. In the traffic control mechanism, the sensor nodes' load is reduced to mitigate the congestion. In the resource control mechanism, the resources are efficiently managed or provide an alternate route. In a hybrid approach, both traffic control and resource control mechanisms are merged [8].

In this context, we propose a resource control based congestion-aware Data gathering (CADA) approach using machine learning for WSNs. This approach provides a dynamic solution for CC in the WSNs. The CADA monitor and determine an efficient congested SNs from the network. Further, it uses the Q-learning mechanism for determining the alternative route for data to the sink. The proposed CADA gives the best result when compared with the existing mechanism through experimental tests. Contribution of proposed CADA is summarized as follows:

- Identify the congestion SNs in the network by computing the buffer occupancy ratio (BOR) and congestion degree.
- Determine the alternative routing path using a Q-learning algorithm to base station/sink.
- Compare the proposed work with the existing CADC and brute force (BF) techniques using various quality metrics such as throughput, energy depletion, packet loss, latency, etc.

The rest of this paper is organized as follows. Section II summarizes the literature study. Section III presents the proposed algorithm in detail. Section IV explains the analysis of simulation results. Finally, we conclude this paper in Section V.

II. Related Work

In this section, we briefly describe various recent and existing CC mechanisms developed for WSNs.

In [12], authors have proposed a congestion-aware data collection (CADC) to improve the accuracy of the data and decrease the estimation error for WSNs and cyber-physical systems. The goal of CADC is to analyze the efficiency of

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

sensor data at the sink with the effect of the CC. To reduce the data distortion, CADC uses the k-means clustering to minimize the transmission rate of WSN. In CADC, the data distinguish with various degrees to decrease the data distortions for improving accuracy. Instead of controlling the resource or transmission rate, authors in [13] introduces a mobile sensor node to alleviate the congestion in WSNs. These mobile nodes create alternative paths towards to BS or use mobile node as a relay node at the congestion places in the network to reduce the burden on the congestion node. This method alleviates congestion efficiently; however, it increases the cost because of the mobile nodes. This method does not explain any way to choose the number of mobile nodes. A Congestion-Adaptive Navigation approach for Sensor networks (CANS) has been proposed in [14] for WSNs. The CANS does not require prior information about the distances and locations in the WSN. It determines the solution based on the message's complexities and time.

In [8], authors have used a PSO (Particle Swarm Optimization) and Single Neuron PID. Initially, the PID helps to queue management, and the PSO perform the online adjustment of the weights to perform self-organizing and self-learning of the nodes in the network. Authors have developed a hybrid multi-objective optimization technique called PSOGSA for regulating the data transmissions between nodes towards the sink. In PSOGSA, the data transmission rates of the nodes controlled depends on the priority of the remaining energy and the bandwidth of SNs. PSOGSA control the number of RTTs by computing the optimal transmission rate. These algorithms mitigate the congestion optimally, but the complexity of the algorithm is high due to the iterative process. In [15], authors have proposed a protocol called RT-CaCC for transport layer with a cache-aware CC for WSNs. The RT-CaCC perform various cache management operations such as cache partitioning to identify the congestion, cache elimination to notify the congestion and cache size (bandwidth) allocation to avoid the congestion. The RT-CaCC protocol is reliable and improves the throughput of the system. However, this method needs to take care of the cross-layer design.

Authors in [16] have proposed a transmission rate control strategy using a support vector machine for WSNs to avoid congestion. This method classifies the congestion nodes efficiently and adjusts the transmission rates to control congestion. But it requires the prior data set to tackle the problem. A distributed adaptive time-division transmission control framework for self-organizing WSNs is developed in [17]to control the congestion. Because of the distributed nature of this approach, the burden on the sensor node is high. An Initial constant congestion window (ICCW) protocol initiate in [18] to avoid the congestion in WSNs. This work also used a mobile element to recharge the sensor nodes in the network. Authors have designed a CC method for WSNs in [19] depends on the data generated by the SN, available energy, and channel availability. A Congestion-sensitive and Load balancing on traffic (CTLS) has been introduced in [20] for WSNs. The CTLS uses a ripple-based search mechanism to control the congestion by detecting it on time. This approach performs congestion detection and then find the alternative solution to alleviate it.

III. Proposed Work

In this section, we discuss the problem definition, congestion node (CN) detection, Q-learning model and congestion-aware data acquisition (CADA) method. In problem definition, we discussed the network model, various parameters used and the goal of the proposed work. This section also describe the process of detecting the CNs and congestion degree. The basic Q-learning and how it can address the issue also described here. Besides, the CADA mainly performs two operations: Identifying CN in the network and Construction of alternative path (CAP) using Q-learning to the BS or sink. Finally, this section derive the complexity of the proposed work.

A. Problem definition

We assume a well connected and un-directed graph $\mathcal{G}(\mathcal{S}, \mathcal{D})$ as a WSN, where $\mathcal{S} = \{\mathcal{S}_0, \mathcal{S}_1, ..., \mathcal{S}_n\}$ indicates the set of SNs and $\mathcal{S}_0$ indicates the BS. The number of SNs in the network is $n = |\mathcal{S}| - 1$. The distance matrix of $\mathcal{S}$ is stored in $\mathcal{D} = \{\delta(\mathcal{S}_i, \mathcal{S}_j) \mid \forall i, j \leq n\}$. Where $\delta(\mathcal{S}_i, \mathcal{S}_j)$ is the distance between two SNs i.e., $\mathcal{S}_i$ and $\mathcal{S}_j$. The limited communication range of the SN is denoted a r. The neighbour node set of each sensor node $\mathcal{S}_i$ is indicated as $NN_i = \{\mathcal{S}_j \mid \delta(\mathcal{S}_i, \mathcal{S}_j) \leq r \ \forall i \neq j \wedge \mathcal{S}_j \in \mathcal{S}\}$. The data collection process of $\mathcal{G}$ from each SN to $\mathcal{S}_0$ (root node) uses the tree topology. The $\mathcal{G}$ partitioned into k trees and each tree indicated as $\mathcal{T}_k$, each $\mathcal{T}_k \subseteq \mathcal{G}$ and $\mathcal{S}_0$ is common root for all $\mathcal{T}_k$. The SNs set of $\mathcal{T}_k$ may varies during the data collection process. The depth of the $\mathcal{T}_k$ is Δ_k i.e., hop count form the child node to $\mathcal{S}_0$. The data transmission in the $\mathcal{T}_k$ is uni-directional towards to the base station. The packet service rate (ϑ) of each sensor node is fixed throughout the simulation. The energy model of the network is adopted from [21].

The PDR is the ratio of the total number of packets received by the sink, i.e., denoted using ($\mathcal{R}$) and the total number of packets transmitted by the SNs, i.e., indicated using (τ) during the time T. We compute the PDR of the system using Eq. (1)

$$PDR = \frac{\mathcal{R}}{\tau} \tag{1}$$

where $\mathcal{R} \leq \tau$, and τ is calculated using the Eq. (2)

$$\tau \cong \sum_{i=1}^{n} P(\mathcal{S}_i) \tag{2}$$

where $P(\mathcal{S}_i)$ is the number of packets collected form the environment by a SN $\mathcal{S}_i$. The network throughput (σ) is the number of packets received by the sink during a unit time T, as shown in Eq. (3).

$$\sigma = \frac{\mathcal{R}}{T} \tag{3}$$

The amount of time taken to receive a packet from SN to the sink is treated as latency ($\mathcal{L}$) of a packet. The $\mathcal{L}$ includes the queuing delay ($\mathcal{L}_q$), radio propagation delay ($\mathcal{L}_r$), signal processing delay ($\mathcal{L}_s$), and transmission delay ($\mathcal{L}_t$). From

these, $\mathcal{L}_r(k) \approx \mathcal{L}_s(k) \leq 1$, so we neglect $\mathcal{L}_r(k)$ and $\mathcal{L}_s(k)$ because of no effect on outcome. The $\mathcal{L}$ of the packet p is calculated as shown in Eq. (4)

$$\mathcal{L}(p) = \begin{cases} \mathcal{L}_q(p) + \mathcal{L}_t(p) & \text{For Successful} \\ (\mathcal{L}_q(p) + \mathcal{L}_t(p)) \times \eta(p) & \text{For re-transmission} \end{cases} \tag{4}$$

where the $\eta(p)$ denotes the number of times, the packet p re-transmitted. The average $\mathcal{L}_a$ of the network is computed as shown in Eq. (5)

$$\mathcal{L}_a = \frac{\sum_{i=1}^{n} \left(\sum_{p=1}^{P(\mathcal{S}_i)} \left(\mathcal{L}(p) \right) \right)}{\mathcal{R}} \tag{5}$$

The algorithm's objective is to minimize the EC and $\mathcal{L}_a$ of the WSNs and improve the PDR and Throughput.

B. Determine Congestion Nodes

The available buffer size of each SN is unique at the beginning, and it varies during the running of the network. It is denoted using $b^t(\mathcal{S}_i)$. The BOR of an SN at time t is computed using Eq. (6).

$$BOR_t(\mathcal{S}_i) = \frac{b^t(\mathcal{S}_i)}{\mathcal{Q}} \tag{6}$$

where $\mathcal{Q}$ indicates the buffer size of an SN. The $BOR_t(S_i)$ returns a value of range [0,1]. However, the BOR is not a perfect parameter to consider for determining the CN, because the data transmissions are continuously competing to occupy the Queue simultaneously. However, BOR_t will helps while constructing the alternate path. So, an alternate parameter uses to identify the CN, i.e., congestion degree (CD) [16], and it is estimated as shown in Eq. (7).

$$\mathcal{C}_d(\mathcal{S}_i) = \begin{cases} 0 & \text{if } \ NN(\mathcal{S}_i) = 0 \\ \frac{T_s \times P_t}{T_a \times (b_t(\mathcal{S}_i) + z)} \times 100 & \text{Otherwise} \end{cases} \tag{7}$$

where $NN(\mathcal{S}_i)$ indicates the number of children nodes for $\mathcal{S}_i$. T_s is the average processing time taken for the data available at $b^t(\mathcal{S}_i)$, P_t denotes the number of packets competing to allocate the buffer of $\mathcal{S}_i$ (it is the sum of packets from all the childs i.e. $NN(\mathcal{S}_i)$) and T_a is the time gap between two data packets to allocate queue of $\mathcal{S}_i$. z is a constant, and its value is either zero or one depending on the $b_t(S_i)$ value. The SN $\mathcal{S}_i$ return the maximum value i.e. $\mathcal{C}_d(S_i) \geq 100$ indicates the node $\mathcal{S}_i$ is congested. When the SN before reaches 100, we consider it a CN, and we start diverting the received nodes from its child nodes.

C. Q-learning Model

The Q-learning is a model-free RL algorithm, and in each iteration, an agent takes a set of states ($\mathcal{S}$) as an input, and it returns a possible action ($\mathcal{A} = \{a_0, a_2, a_3, a_4\}$). The four actions in the sequence are choosing the new node for data transmission, Send as it is like last packet transmission path, hold in the queue for some more time, and drop the packet. The action selection a_i^t is depending on the policy (π) based on the next state election. The agent considers the policy (π) to determine the best action a_i^t by using an adjustable parameter θ and current state $\mathcal{S}_i^t$. It is denoted, as shown in Eq. (8).

$$\pi_\theta(s_i^t, a_i^t) = arg\max_a \ Q(s, a) \tag{8}$$

Before moving to the next state, the agent receives a reinforcement or reward ($R_t(a_t)$). The reward function in Q-learning is important, and it decides the agent's learning capabilities. The Goal of Q-learning is to maximize the sum of reward while identifying the sequence of actions. In this work, the reward function is computed, as shown in Eq. (9).

$$R_t(\mathcal{S}_i, a_i) = \sum_{i=t}^{\inf} \gamma^{(i-t)} \varphi_i(\mathcal{S}_i, a_i) \tag{9}$$

where φ_i is computed as shown in Eq. (10) and $\gamma \in (0, 1]$.

$$\varphi_i(\mathcal{S}_i, a_i) = \left(\frac{\sum_{i-1}^{NN(\mathcal{S}_i)} BOR_t(\mathcal{S}_i)}{|NN(\mathcal{S}_i)|} \right) \tag{10}$$

The Q-value function can be updated based on the Eq. (11)

$$\begin{aligned} Q_{t+1}(\mathcal{S}_i, a_i) = (1-\alpha) Q_t(\mathcal{S}_i, a_i) + \alpha \times R_t(\mathcal{S}_i, a_i) \\ + \alpha \times \Upsilon \left(\max_a Q^{'}(\mathcal{S}_i^{'}, a_i^{'}) - Q_t(\mathcal{S}_i, a_i) \right) \end{aligned} \tag{11}$$

where Υ is the discount factor, α is the learning rate, $Q^{'}(\mathcal{S}_i^{'}, a_i^{'})$ indicates the maximum anticipated next reward given the new state $\mathcal{S}^{'}$ and all feasible actions at $\mathcal{S}^{'}$. The $Q_t(\mathcal{S}_i, a_i)$ is the current Q-values.

D. Congestion-aware Data Acquisition

The proposed Congestion-aware Data Acquisition (ACDA) algorithm iterates three steps i.e., Spanning tree contraction, congestion node detection, and an alternate route construction. Initially, the algorithm contracts the spanning tree from the child node to the sink using prim's algorithm. The $\mathcal{C}_d$ is computed at each of the $\mathcal{S}_i$ to determine the CNs from the network using Eq. (7). Once the CN is identified, we need to divert the downstream node data to another optimal upstream node by avoiding the CN, which is within a single-hop, as shown in Fig. 1. To achieve this, the proposed CADA algorithm introduced using Q-learning.

The reward matrix R is constructed based on r of the SNs with the dimensions of $|\mathcal{S}| \times |\mathcal{A}|$. The Eq. (12) represents the initial R matrix. Each cell values in the matrix represent the corresponding state's $\mathcal{S}_i$ reward value to determine the next node.

$$R = \begin{array}{c} \\ \mathcal{S}_0 \\ \mathcal{S}_1 \\ \vdots \\ \mathcal{S}_n \end{array} \begin{array}{c} \begin{array}{ccccc} a_1 & a_2 & ... & ... & a_n \end{array} \\ \begin{bmatrix} -1 & -1 & 0 & ... & 100 \\ -1 & 0 & 0 & ... & 100 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & 0 & 0 & ... & -1 \end{bmatrix} \end{array} \tag{12}$$

There are three possible values in the Eq. (12), in which -1 indicates there is no communication between the nodes

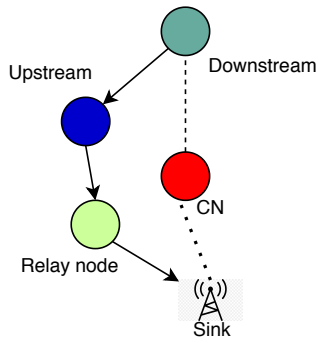


Fig. 1: Alternate Routing Path for CADA

because of less r or CN. The cell value zero (0) indicates the SNs are in r but not transmitting their data. The cell value < 0 indicates the nodes are in r and transmitting their data among the nodes, and it also may be the reward value. Some of the SN's value maybe change to -1 if the node is converted as CN. The initial Q matrix is represented in Eq. (13). The dimension of the Q matrix is similar to R matrix it is $|\mathcal{S}| \times |\mathcal{A}|$.

$$Q = \begin{array}{c} \\ \mathcal{S}_0 \\ \mathcal{S}_1 \\ \vdots \\ \mathcal{S}_n \end{array} \begin{array}{c} \begin{array}{ccccc} a_1 & a_2 & ... & ... & a_n \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & ... & 0 \\ 0 & 0 & 0 & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 0 \end{bmatrix} \end{array} \tag{13}$$

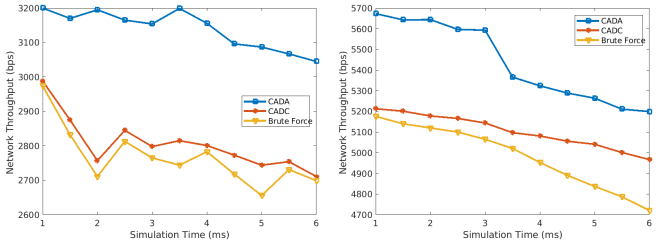
The base station starts collecting all the network information by sending a beacon signal to each SN. Once the Q and R matrices are assigned, the base station starts training and running process. In the training process, the sink node iterative the following steps:

1) Select the *current state* ($\mathcal{S}_i$) randomly from the $\mathcal{S}$.
2) Identify all the possible *next nodes* to whom the data can be transmitted.
3) Select one *next node* randomly from the possible next nodes list
4) Use the Eq. (11) to update the R and Q matrices in the Eq. (12) and Eq. (13).

The training process continuously updates the Q matrix. During the data collection phase, irrespective of the training process, the running process also performs simultaneously.

1) Randomly choose the *initial state* and store in a
2) Identify the state $\mathcal{S}_i$ based on the maximum value of the row a.
3) if $\mathcal{S}_i$ is not the base station, the statements (1) & (2) will repeat.

However, the proposed algorithm performs better congestion control over the WSN data transmissions; there are still some limitations. The CADA may increase the latency compare with the regular data transmissions. It is happening because of choosing the longer path to avoid the CNs. This algorithm is centralized, and it runs on the BS. The energy consumption (EC) of the SN will increase if it runs a decentralized manner.



(a) (b)

Fig. 2: Network Throughput of (a) WSN#1 (b) WSN#2

E. Complexity Analysis

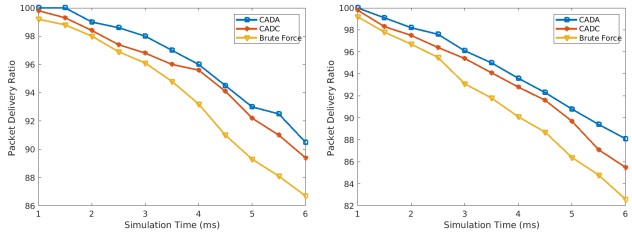
The proposed CADA takes the time to compute the spanning tree, CN detection, and an alternate path determination. The computational time for spanning tree construction using prim's algorithm is $O(E+VlogV)$, where E indicates the number of edges, and V indicates the number of vertices. The time taken to identify the $\mathcal{C}_d$ of each SN in the WSN is $O(n)$, where n indicates the number of SNs. The worst-case computational complexity to determine the alternate path using Q-learning is $O(ss^3)$ (Assuming no duplicate actions), where ss is the state space [22]. The relation between state space and $\mathcal{S}$ is $ss \approx n$. So, The time complexity of Q-learning may consider as $O(n^3)$. CADA's overall computational complexity can be written as $O(E+VlogV)+O(n)+O(n^3)$ and asymptotically the time complexity is $O(n^3)$.

IV. Experimental Results

The proposed and existing mechanisms are implemented using Python simulator (python v3.7.4). The buffer size of the SN is considered as 512 KB. The EC for transmission and receive are considered as 17.4 mA and 19.4 mA, respectively. The packet size various between 20 Bytes to 127 Bytes. We consider two network scenarios i.e. WSN#1 and WSN#2. In WSN#1, we consider 100 to 200 SNs in 200 sq. m whereas in WSN#2 500 to 1000 sq. m with 1000 SNs. In both scenarios the nodes are deployed randomly. The r value is considered as 15 m to 35 m. The initial energy of each SN is consider as 100 J. The data transmission rate considered as 40 Kbps. The proposed CADA is compared with existing CADC [12], and brute force (BF) algorithms with various performance metric shown below from Section IV-A to IV-F.

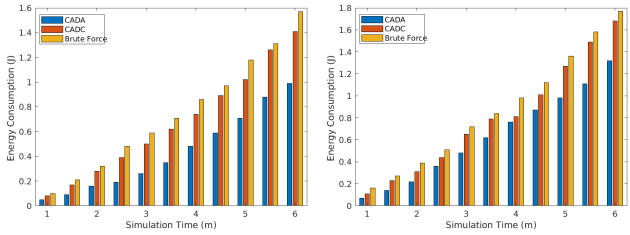
A. Network Throughput

The network throughput (σ) is determined as a unit time, the number of packets received by the sink. It is calculated using the Eq. (3). The σ is inversely proportional to the $\mathcal{C}_d$. The maximum σ indicates the better performance of the network and vice versa. Fig. 2 demonstrates the improved throughput during the simulation runs. From Fig. 2(a), the throughput of the network is improved over the existing CADC method around 16%-18%. The improved throughput in the BF algorithm is ≈18%-21%. Fig. 2(b) shows that the the performance improvement of scenario two of both existing and proposed works. In CADC the performance improved



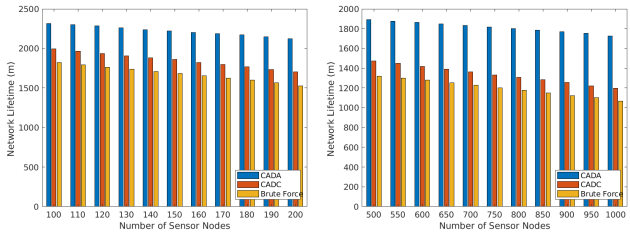
(a) (b)

Fig. 3: Packet Delivery Ratio of (a) WSN#1 (b) WSN#2



(a) (b)

Fig. 4: Average EC of (a) WSN#1 (b) WSN#2



(a) (b)

Fig. 5: Network Lifetime of (a) WSN#1 (b) WSN#2

approximately 9%-12% and in BF the performance improved approximately 11%-15%. During the simulation, in both the scenarios, we consider 200 SNs for WSN#1 and 1000 SNs in WSN#2. The performance improvement is achieved because of the best alternative path and efficient identification of the CNs.

B. Congestion Fairness Index

The Congestion Fairness Index (CFI) shows the "equal share of the bottleneck of congestion among the WSN" [23]. The CFI and $\mathcal{C}_d$ are inversely proportional to each other. We consider the computation of CFI according to the [21]. The fairness of CADA in WSN#1 is ≈0.99828 when considering the 200 SNs. In WSN#2 scenario, the CFI is ≈0.98119. The CFI of existing approaches for WSN#1 is 0.956625 for CADC and 0.8789 for the BF approach. Similarly, the CFI of CADC and BF for WSN#2 are ≈0.94834 and ≈0.8598, respectively. Higher the fairness index indicates less congestion balance over the WSN.

C. Packet Delivery Ratio

The PDR is computed as shown in Eq. (1). It is inversely proportional to the $\mathcal{C}_d$ and directly proportional to σ. Fig. 3, shows the comparison of PDR of proposed CADA with the existing approaches. In this, the results are collected with various simulation time changes. From Fig. 3(a), the PDR is improved around 8%-10% in CADA compared it with the CADC and 12%-15% when compared with the BF approach. From Fig. 3(b), the PDR is improved around 7%-10% in the second scenarios of CADC and 10%-14% when compared it with the BF approach. Here also, we consider 200 SNs for WSN#1 and 1000 SNs in WSN#2. In CADA, still, the packet loss is happening when no alternative path available to data transmissions.

D. Average Energy Consumption

The EC is one of the most important parameters for WSNs because it is operating with a limited battery. The average EC (μ_e) and $\mathcal{C}_d$ are inversely proportional to each other. The average energy consumption (AEC) of the network is computed as shown in Eq. (14).

$$\mu_e = \frac{\sum_{i=1}^{n} \mathcal{E}(\mathcal{S}_i)}{n} \tag{14}$$

where $\mathcal{E}(\mathcal{S}_i)$ is the energy consummation of each SN ($\mathcal{S}_i$) and n is the number of SNs. Fig. 4 shows the AEC of the CADA, CADC, and BF algorithms. From Fig. 4(a), we observe that the AEC of the proposed CADA algorithm improved ≈7%-12% compared with the CADC and 11%-17% compare with the BF approach when considered the simulation time. Fig. 4(b), shows the improvement of the CADA over the CADC is ≈6-11%. Similarly, the AEC varies in the BF ≈9%-14%. The AEC of the proposed work is improved because of minimizing the re-transmissions of the data packets.

E. Network Lifetime

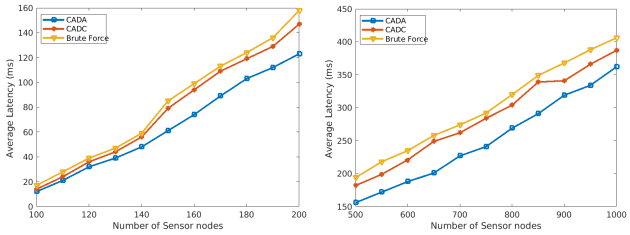
It is determined as the time the network is operational. The comparisons of the network lifetime (NL) for the proposed CADA, and existing CADC and the BF are shown in Fig. 5. Minimizing the re-transmissions with optimal routing path reduces the AEC, and minimum EC maximizes the NL. The NL of the proposed and existing algorithms are calculated using Eq. (15)

$$NL = \frac{\mathcal{E}_0}{\Gamma} \tag{15}$$

where $\mathcal{E}_0$ indicates the initial energies of the SNs and Γ denotes the maximum energy consumed by any of the SN in the network and it is calculated as shown in Eq. (16). Mostly the more energy consumed by the nodes which are affected congestion.

$$\Gamma = \max_{\mathcal{S}_i \in \mathcal{S}} (\mathcal{E}(\mathcal{S}_i)) \tag{16}$$

Fig. 5(a), we observe the NL improved ≈10%-16% compare with the CADC and ≈12%-17% compare with the BF. The improvement of NL in scenario two is shown in Fig. 5(b). From Fig. 5(b), the growth of the NL compared with CADC and BF is ≈ 9%-14% and 10%-15%, respectively. The NL



(a) (b)

Fig. 6: Average Latency of (a) WSN#1 (b) WSN#2

growth is achieved because of the efficient congestion handling mechanism effected by the proposed work using Q-learning.

F. Average Latency

The average time taken to receive a packet from its source to the base station is considered Average Latency. It is inversely proportional to σ, PDR, and directly proportional to $\mathcal{C}_d$. Fig. 6 shows the comparison of the latency among the proposed and existing works. The latency is computed using Eq. (4), and the average latency is computed as shown in Eq. (5).

From Fig. 6(a), the latency of CADA is reduced around 8%-13% in the CADC approach 10%-16% in the BF method by varies SNs between 100-200. Fig. 6(b) shows the latency variations in the proposed and existing, and the number of SNs changing in this scenario is between 500 to 100 nodes. The improvement of the latency minimization compare with the CADC is ≈6%-11%, and the BF is ≈9%-14% in the scenario two. The latency of the proposed and existing methods are still a challenging because of the retransmission controls and longer smooth path.

V. Conclusion

Congestion in the wireless sensor network is a challenging issue that affects various QoS metrics of the WSNs. It mainly occurs because of the uncontrollable traffic in the WSNs. The congestion can be handle by controlling the traffic, efficient resource management, or both. A resource control based congestion-aware data acquisition (CADA) mechanism for wireless sensor networks is proposed in this paper. The CADA efficiently identifies the network's congestion nodes and applies the alternative routing path to the base station. In this way, the proposed algorithm efficiently mitigating the congestion compared with the existing works. The simulation results show that the proposed method outperforms the existing techniques in throughput, packet loss, energy consumption, and latency. This work can be extended in the future to predict the congestion before it occurs in the WSNs and take an alternative decision to mitigate the congestion to avoid the latency.

References

[1] D. P. Kumar, T. Amgoth, and C. S. R. Annavarapu, "Machine learning algorithms for wireless sensor networks: A survey," *Information Fusion*, vol. 49, pp. 1–25, 2019.

[2] C. Sergiou, P. Antoniou, and V. Vassiliou, "A comprehensive survey of congestion control protocols in wireless sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 1839–1859, 2014.

[3] G. Künzel, L. S. Indrusiak, and C. E. Pereira, "Latency and lifetime enhancements in industrial wireless sensor networks: A q-learning approach for graph routing," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5617–5625, 2019.

[4] A. Mukherjee, S. Misra, V. S. P. Chandra, and M. S. Obaidat, "Resource-optimized multiarmed bandit-based offload path selection in edge uav swarms," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4889–4896, 2018.

[5] H. B. Liaqat, A. Ali, J. Qadir, A. K. Bashir, M. Bilal, and F. Majeed, "Socially-aware congestion control in ad-hoc networks: Current status and the way forward," *Future Generation Computer Systems*, 2019.

[6] M. A. Jan, S. R. U. Jan, M. Alam, A. Akhunzada, and I. U. Rahman, "A comprehensive analysis of congestion control protocols in wireless sensor networks," *Mobile Networks and Applications*, vol. 23, no. 3, pp. 456–468, 2018.

[7] V. Srivastava, S. Tripathi, K. Singh *et al.*, "Energy efficient optimized rate based congestion control routing in wireless sensor network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 3, pp. 1325–1338, 2020.

[8] X. Yang, X. Chen, R. Xia, and Z. Qian, "Wireless sensor network congestion control based on standard particle swarm optimization and single neuron PID," *Sensors*, vol. 18, no. 4, p. 1265, 2018.

[9] C. Sergiou, V. Vassiliou, and A. Paphitis, "Hierarchical tree alternative path (htap) algorithm for congestion control in wireless sensor networks," *Ad hoc networks*, vol. 11, no. 1, pp. 257–272, 2013.

[10] K. Xiao, S. Mao, and J. K. Tugnait, "TCP-Drinc: Smart congestion control based on deep reinforcement learning," *IEEE Access*, vol. 7, pp. 11 892–11 904, 2019.

[11] D. Pandey and V. Kushwaha, "An exploratory study of congestion control techniques in wireless sensor networks," *Computer Communications*, 2020.

[12] Y. Zhuang, L. Yu, H. Shen, W. Kolodzey, N. Iri, G. Caulfield, and S. He, "Data collection with accuracy-aware congestion control in sensor networks," *IEEE Transactions on Mobile Computing*, vol. 18, no. 5, pp. 1068–1082, 2019.

[13] A. Nicolaou, N. Temene, C. Sergiou, C. Georgiou, and V. Vassiliou, "Utilizing mobile nodes for congestion control in wireless sensor networks," *arXiv preprint arXiv:1903.08989*, 2019.

[14] C. Wang, H. Lin, and H. Jiang, "CANS: Towards congestion-adaptive and small stretch emergency navigation with wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 15, no. 5, pp. 1077–1089, 2016.

[15] M. I. Alipio and N. M. C. Tiglao, "RT-CaCC: A reliable transport with cache-aware congestion control protocol in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, pp. 4607–4619, 2018.

[16] M. Gholipour, A. T. Haghighat, and M. R. Meybodi, "Hop-by-Hop congestion avoidance in wireless sensor networks based on genetic support vector machine," *Neurocomputing*, vol. 223, pp. 63–76, 2017.

[17] A. Kamimura and K. Tomita, "A self-organizing network coordination framework enabling collision-free and congestion-less wireless sensor networks," *Journal of Network and Computer Applications*, vol. 93, pp. 228–244, 2017.

[18] N. Aslam, K. Xia, A. Ali, and S. Ullah, "Adaptive TCP-ICCW congestion control mechanism for QoS in renewable wireless sensor networks," *IEEE Sensors Letters*, vol. 1, no. 6, pp. 1–4, 2017.

[19] P. D. Godoy, R. L. Cayssials, and C. G. G. Garino, "Communication channel occupation and congestion in wireless sensor networks," *Computers & Electrical Engineering*, vol. 72, pp. 846–858, 2018.

[20] O. Chughtai, N. Badruddin, A. Awang, and M. Rehan, "Congestion-aware and traffic load balancing scheme for routing in WSNs," *Telecommunication Systems*, vol. 63, no. 4, pp. 481–504, 2016.

[21] P. K. Donta, B. S. P. Rao, T. Amgoth, C. S. R. Annavarapu, and S. Swain, "Data collection and path determination strategies for mobile sink in 3D WSNs," *IEEE Sensors Journal*, vol. 20, no. 4, pp. 2224–2233, 2020.

[22] S. D. Whitehead, "A complexity analysis of cooperative mechanisms in reinforcement learning," in *AAAI*, 1991, pp. 607–613.

[23] T. HoBfeld, L. Skorin-Kapov, P. E. Heegaard, and M. Varela, "Definition of QoE fairness in shared systems," *IEEE Communications Letters*, vol. 21, no. 1, pp. 184–187, 2017.

Novel Multi-logic gates using Quantum Dot Cellular Automata with energy dissipation analysis

Avinashkumar M
Department of Electrical and Electronics Engineering Birla Institute of Technology and Science Pilani ,Dubai Campus,
Dubai, UAE
avimkumar16@gmail.com

Anuj Borkute
Department of Electrical and Electronics Engineering Birla Institute of Technology and Science Pilani ,Dubai Campus,
Dubai, UAE
anujb996@gmail.com

Dr. Nilesh Goel
Department of Electrical and Electronics Engineering Birla Institute of Technology and Science Pilani ,Dubai Campus,
Dubai, UAE goel.nilesh@gmail.com

Abstract— Quantum dot Cellular Automata (QCA) poses novel computing paradigm to design Nano electronic circuits with favorable proximity intended towards high speed, energy efficient and compact devices. Hence this emerges as a viable beyond CMOS (complementary metal oxide semiconductor) Technology. Many literatures are based on 3 input-1 output majority gate logic, which primarily focuses to reduce the number of cells in single individual logic gates utilized. In this work we are extending our proposed multi-logic gates with different configurations to produce multiple logic functions with the same inputs of individual logic gates and also performing energy dissipation analysis. These gates are very useful since this enables the design of efficient smaller circuits. Proposed multi-logic design consumes less area than that of individual gates. The design is simulated in QCA Designer 2.0 and found in accordance with the logic. An analysis to explore the power dissipated by multi-logic gates is carried out by QCA Designer-E. The area and power efficiency of these multi-logic gates are compared with the previous best designed individual gates. Superior results are obtained and reported here.

Keywords—QCA gates, Multi logic gates, QCA Designer, Nanotechnology, Cellular Automata

I. INTRODUCTION

Rapid and dynamic CMOS scaling eventually approaches the atomic size which faces many fundamental challenges such as excessive Gate leakage current, exponentially increasing source to drain sub-threshold leakage current, gate stack reliability and channel mobility degradation from increasing electric field, rising dynamic power dissipation (CV2f) from non-scaled supply voltages, band to-band tunneling leakage at high body doping levels, device to device variation from random dopant fluctuation effects, and high source-drain access resistance from scaled contact areas limiting on-current[1-3]. This seems major threat for further shrinking of devices.

Hence to keep up with the Moore's law, innovation in computing trend is the need of the hour. Many architectures have been proposed for future Nano electronics such as single electron transistor, carbon nanotubes, Quantum computing & Quantum dot cellular automata [4, 5]. Among these QCA emerges as a promising alternative providing logical constructs at Nano scale level [6]. Here structure and working of QCA circuits are completely different from the traditional CMOS circuits and codes binary information 0 and 1 as electronic configuration within its Quantum cells [7]. Each cell is a square block with four Quantum dots at each corner where two electrons are locked in a diametrically opposite manner. Electrons within Quantum cell play an important role in the flow of binary information from input to output. Switching takes place due to columbic interaction between the cells hence does not rely on the voltage level as the traditional CMOS logic [8, 9]. This favors the desired computation logic, developing into low power Nano devices.

QCA circuits depend on inverter and majority (three inputs-one output) logic [10-14]. These two elements form the basis of all individual logic gates. Here in this paper we propose the design of multi-logic gates which can produce several multiple function outputs with the same inputs as the individual gates, in short, we can design intricate circuits using these novel multi logic gates with optimization in area consumed and delay. All these enhanced designs are compared with the previous best designs using QCA cells which show that these designs consume less area and power compatible with the individual logic gates. This work is organized in 5 sections. The section 2 brings out the fundamentals of QCA. The proposed model and its configuration are presented in section 3. The desired result and the power consumed by these gates are explained in section 4. The paper is concluded with the future possibilities in section 5.

II. QCA FUNDAMENTALS

A quantum cell is a major building block of any QCA design [15]. Each square block cell consists of four quantum dots which are coupled to each other due to columbic force and there is a barrier between each of these dots which is capacitive in nature. Figure 1(a) shows basic QCA cell with capacitive barriers [16, 17]. Quantum dots exhibit a phenomenon called columbic exclusion, when two cells are kept together which leads to an effect called quantum confinement and it gives two stable states for each cell which then can be used in cellular array. When multiple such cells put together leads to form cellular automata called as Quantum Cellular Automata and the interaction between cells can be made to implement logical functions [18].

Electrons can be injected into quantum cells but because polarization value can be found out with equation 1. The information between various cells is confined with the help of four different clocks.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Figure 2 shows all clock phases employed in QCA. Electrons move in different quantum dots through junction tunneling which is controlled by the capacitive barrier between the neighboring cells [20, 21] which can be raised high or low by the adiabatic clock cell given. Each clock has four phases as described in table 1. Potential barrier becomes high in switch phase and cell gets polarized. Hold phase keep potential barrier high and cell polarized. In release and relax phase potential barrier becomes low and remain low respectively and cell is un-polarized. Switching takes place when clock goes high, cells gets polarized and the information gets confined.

$$p = \frac{P2+P4-(P1+P3)}{P1+P2+P3+P4} \quad (1)$$

Because of cellular automata each cell interacts with each other resulting in transmission of data from input to output [22]. In QCA circuit the polarized charge with columbic repulsion carries information and not with current, so these circuits have low energy with high speed. Figure 3 shows a QCA wire demonstrating flow to information from input to output.

TABLE I. QCA CELL POLARIZATION AT DIFFERENT CLOCK PHASES

Clock Phase	*Potential Barrier*	*Cell polarization state*
Switch	Low to High	Polarized
Hold	Held High	Polarized
Release	Low	Un-polarized
Relax	Remain Lowered	Un-polarized

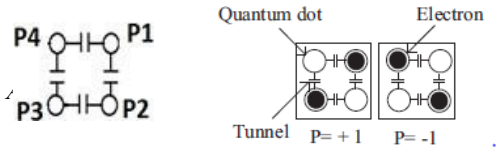


Fig. 1. (a) Basic QCA Cell. (b) QCA Cells with different electron position

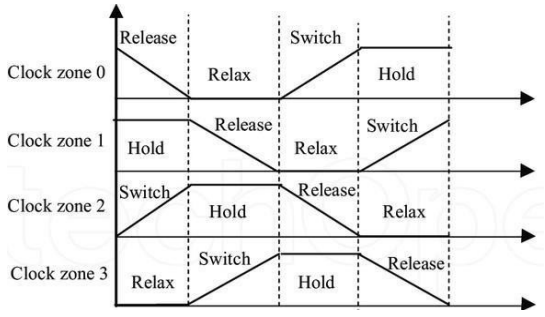


Fig. 2. Clock zones in QCA

A. Basic QCA gates

Logic functions are implemented by connecting QCA cells with each other with some specific design. We can fix the polarization of particular QCA cell to get desired logic [23-24]. Inverter & majority gate logics are the two important elements to design QCA circuits. Figure 4 shows different basic gates designed with QCA cell. (a) is inverter design having two arrays of cells at 45o to each other, (b) is a three- input majority gate, at the output Y of which we get one of the inputs from A, B, C according to majority of polarization, equation 2 denotes majority gate equation. From majority gate AND and OR basic digital gates can be implemented by giving polarization 0 and 1 to input C as shown in (c) and (d) respectively. Equation 3 and 4 denotes equation of AND and OR gate from majority gate equation

$$Y(A, B, C) = AB+BC+AC \quad (2)$$

$$Y(A, B, 0) = A.B \quad (3)$$

$$Y(A, B, 1) = A+B \quad (4)$$

All digital logic functions can be implemented using inverter and majority gate [25]. Majority gate acts as universal gate to form different logic functions. Figure 5 demonstrate by combining majority gate and inverter, NAND [23-27], NOR [23-27], XOR [23-27] and XNOR [23-27] can be designed.

Fig. 3. QCA Wire

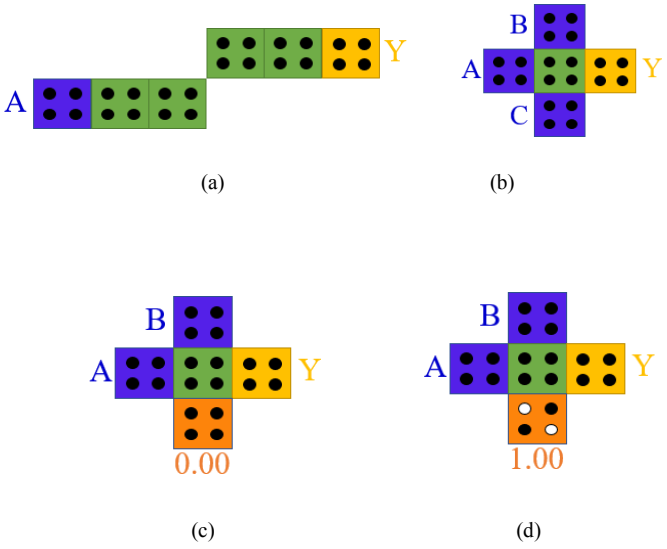


Fig. 4. (a) Inverter (b) Majority gate (c) AND Gate (d) OR Gate

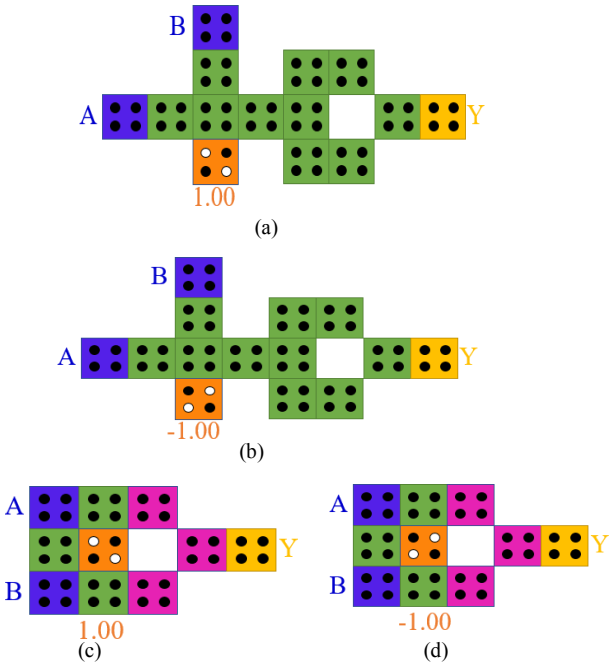


Fig. 5. (a) NOR Gate (b) NAND Gate (c) EXOR Gate (d) EXNOR Gate

III. PROPOSED MULTI-LOGIC GATES AND QCA IMPLEMENTATION

Circuits that implement many functions can be combined into a single compact circuit with the help of multi logic gates. Complex circuits such as an ALU make use of these kinds of multi logic gates in CMOS design, hence designing these fundamental gates becomes one of the important areas of research in QCA. Yet this path is not much explored till now. Traditional Boolean gates have 2 inputs and 1 output as shown in figure 6(a) and figure 6(b). Our proposed design is like a combined logic as shown in figure 6(c). Combined logic design (multi-logic gates) generates two different logic outputs (F and F') from two same inputs (A and B) as shown in figure 6(c). In QCA design we try to implement architecture like figure 6(c). Proposed QCA designs has less number of cells, latency, less area and gate count compared with the individual separate logic design using QCA.

Polarization is fixing arrangement of two electrons in a cell such that particular cell always represents binary '1' or binary '0'. Polarization P = 1 represents binary '1' and P = '-1' or '0' represents binary '0'. Also we can provide any value between 0 to 1 and 0 to -1 as a polarization P to a particular cell which is called partial polarization and according to that electrons arrange themselves in that cell with particular energy. Partially polarized cell interacts with neighboring cells differently according to polarization value assign to it.

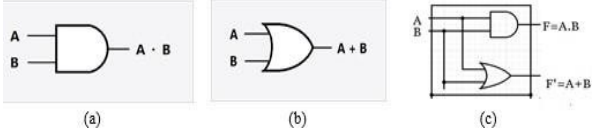


Fig. 6. (a) Individual AND Gate (b) Individual OR Gate (c) Combined multi-gate logic example

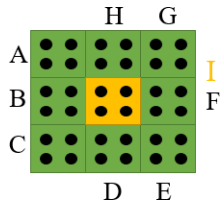


Fig. 7. Neighborhood cell

So we use partial polarization technique and 8 neighborhood concept in the design of these multi-logic QCA cells. According to 8 neighborhood/Moore neighborhood, the finite neighborhood of a QCA cell N = {N1, N2, N3…Nr} has effect on the current state of cell. For example in Figure 7 the current state of QCA cell I depends on previous state of four neighbors (B, D, F, H) and cells (A, C, G, E) along the diagonal direction.

A. AND-OR Multi Logic Design

Proposed AND-OR multi logic QCA gate design is shown in figure 8(a). Cell at center is polarized with fixed value -0.40 and cell with output Y' is polarized with fixed value 0.40. When inputs given, we get AND & OR outputs at Y & Y' respectively. A QCA cell can be rotated and if we rotate both the output cells Y & Y' by 90° in same design, we obtain NAND-NOR logic. Figure 8(c) represents OR-NOR and AND-NAND design. Here 2 cell inverter is connected to output Y' of AND-OR design and only center cell is polarized. Table 2 explains that when polarization value is fixed as 0.40, OR output can be obtained at Y and NOR output at Y'. Similarly when polarization value is fixed as - 0.40, we get AND function at output Y and NAND function at output Y'.

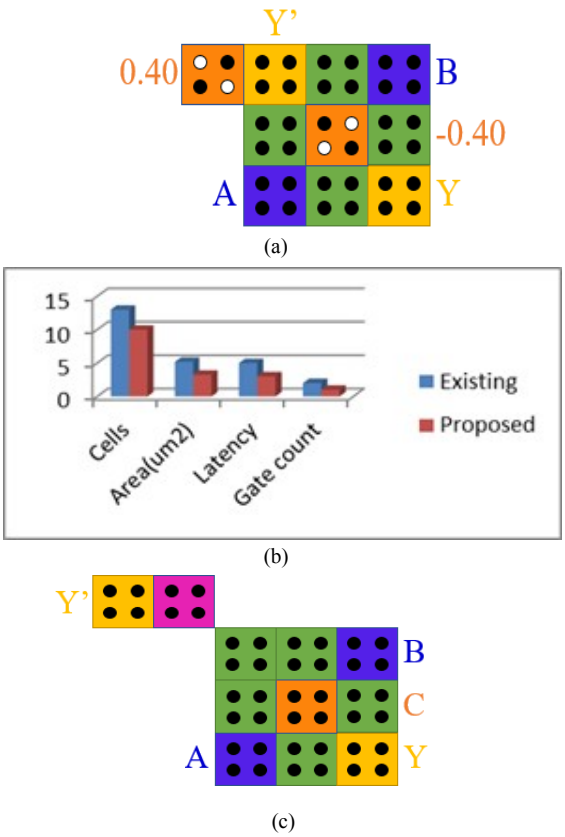


Fig. 8. (a) AND-OR gate design (b) Comparison of Proposed and existing design [28](c) OR-NAND Design and AND-NOR Design

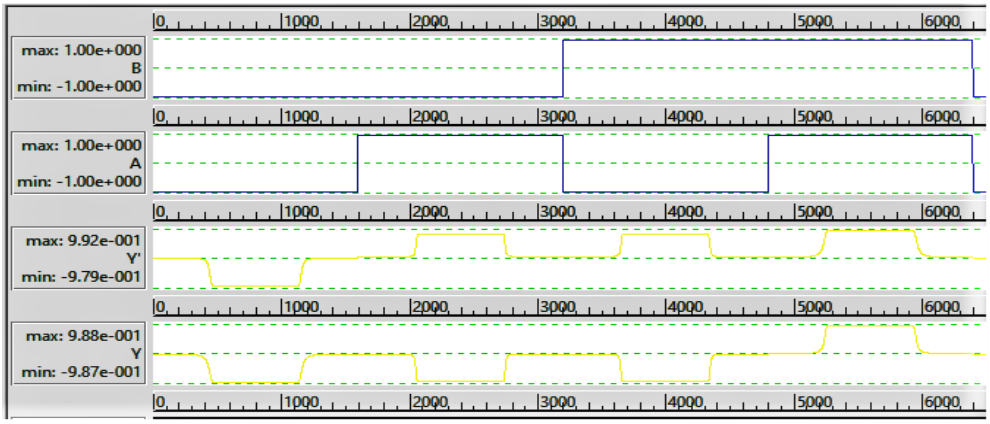


Fig. 9. Simulation result of AND-OR gate

Table II. OR-NOR and AND-NAND Design

Input			Output	
A	***B***	***Polarization Value C***	***Y***	***Y'***
A	B	0.40	A+B	(A + B)'
A	B	-0.40	A.B	(AB)'

The simulation result of proposed AND-OR gate is shown in the Figure 9. Figure 8(b) shows a comparison between proposed multi-logic AND-OR design and existing AND-OR design [28] in terms of no of cells, area consumed, latency for output, and the gate count required to design AND-OR gate. Proposed design uses fewer cells for the same result consuming less area. Also it is faster as the gate count is less, so delay to reach output for given input is less.

B. OR-NAND and AND-NOR Multi Logic Design.

Here the center cell can be provided with fixed polarization of +0.40 and -0.40. 2-cell invertors is connected to output of AND-OR gate design with corner cell confined to clock 1, to obtain OR –NAND and AND-NOR design as shown in figure 10 below. Table 3 explains that when polarization value is fixed as 0.40, OR output can be obtained at Y and NAND at Y'. Similarly when polarization value is fixed as -0.40 we get AND function at output Y and NOR at output Y'.

C. EXOR-OR & EXNOR-AND Multi Logic Design

Figure 11 represents the design of EXOR-OR and EXNOR-AND gate. Here cell with input C is polarized cell. Table 4 explains that when polarization value is fixed as 1, at output Y we get EXOR function of inputs and at output Y' we get OR function of given inputs. Similarly if polarization value is fixed as -1 at output Y, EXNOR function is obtained and at output Y', AND function can be obtained.

D. EXOR-AND & EXNOR-OR Multi Logic Design

EXOR-AND and EXNOR-OR gate is shown in Figure 12 below. Here cell with input C is polarized cell and fixed polarization of 1 and -1 can be provided to it.

Table 5 explains that when polarization value is fixed as 1 at output Y, EXOR And at output Y', AND function can be obtained. Similarly if polarization value is -1, then EXNOR and OR function of given inputs can be obtained at Y and Y' respectively.

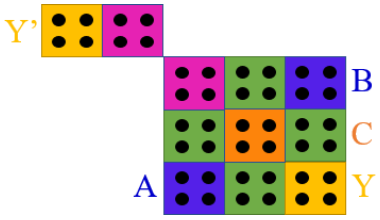


Fig. 10. OR-NAND Design and AND-NOR Design

TABLE III. OR-NAND AND AND-NOR DESIGN

Input			Output	
A	***B***	***Polarization Value C***	***Y***	***Y'***
A	B	0.40	A+B	(AB)'
A	B	-0.40	A.B	(A+B)'

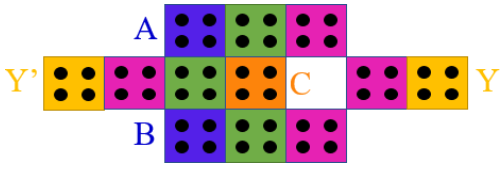


Fig. 11. EXOR-OR gate Design and EXNOR-AND gate design

TABLE IV. EXOR-OR AND EXNOR-AND DESIGN

Input			Output	
A	***B***	***Polarization Value C***	***Y***	***Y'***
A	B	1	A EXOR B	A + B
A	B	-1	A EXNOR B	AB

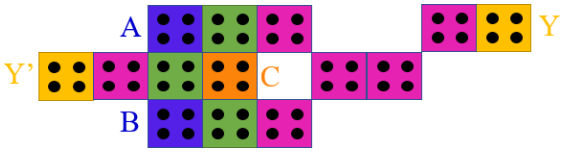


Fig. 12. EXOR-AND and EXNOR-OR gate design

TABLE V. EXOR-AND AND EXNOR-OR DESIGN

Input			Output	
A	*B*	*Polarization Value C*	*Y*	*Y'*
A	B	1	A EXNOR B	A +B
A	B	-1	A EXOR B	AB

E. EXOR-EXNOR Multi Logic Design

To get the EXOR and EXNOR from the single design, an inverter can be used with the existing EXOR design. So if we extend existing design by adding 2 cells inverter at output of EXOR gate we obtain EXNOR output. At output F we get EXOR output and at F' we get EXNOR output.

Figure 13(a) shows a proposed EXOR-EXNOR gate design and Figure 14 shows its simulation result. In EXOR- EXNOR design fix polarization of -1 is provided at one end opposite to input B. if we add another that is third input C at place of polarization, we get three input EXOR-EXNOR design. Figure 13(b) shows three input EXOR-EXNOR design.

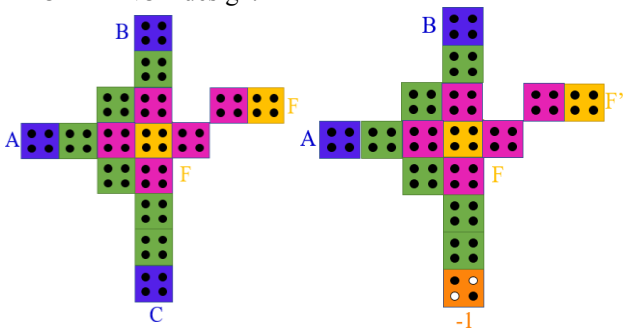


Fig. 13. (a) Proposed EXOR-EXNOR design (b) Three input EXOR-EXNOR design

IV. RESULTS AND DISCUSSIONS

The proposed multi-logic QCA designs has been verified and designed utilizing software called QCA Designer version 2.0.3 which is provided with the default parameters of coherence vector and bi-stable engines [30- 32]. The identical results from both engines prove the accurateness of proposed model. The results of proposed designs are shown in Table 6 which demonstrates the latency, number of cells, area consumed of these designs and it is compared with individual gates. Figure 15 shows graphical representation of area occupied in 10-2 µm2 by multi-logic design and combination of individual gates. Except EXOR-EXNOR multi-logic design area occupied by proposed multi-logic designs is less than that of combination of individual gates and number of cells used are also reduced in multi-logic design in comparison to sum of individual gates.

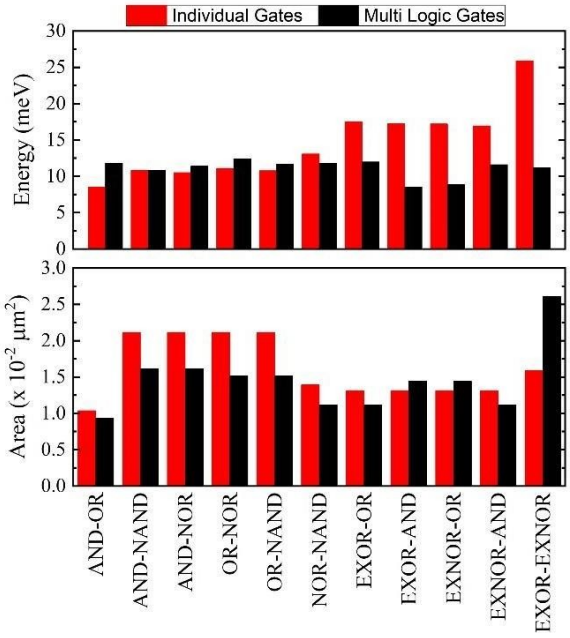


Fig. 15. Power dissipated by individual and proposed Gates Multi-logic gates

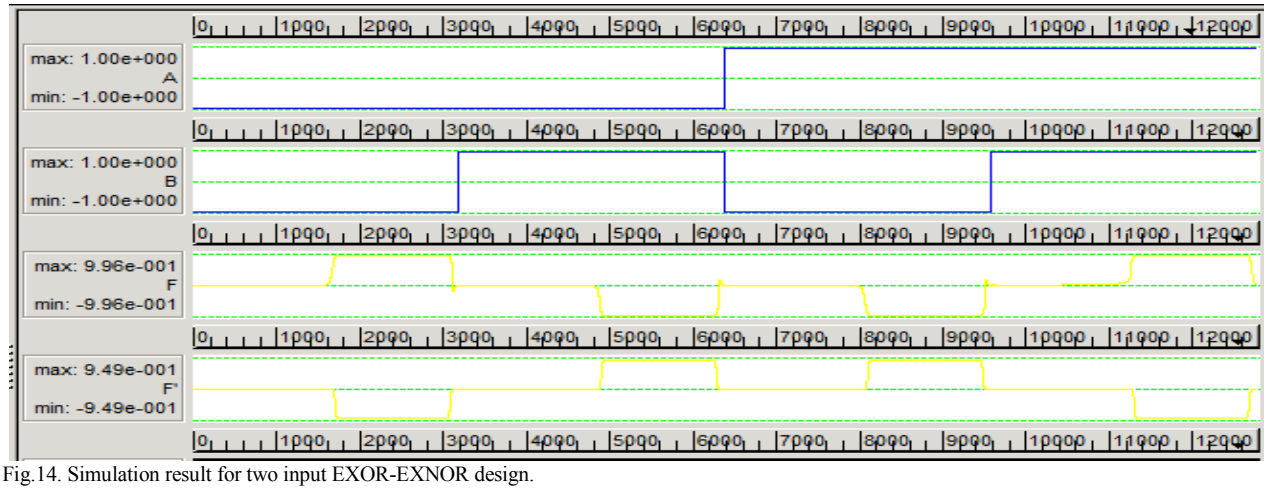


Fig.14. Simulation result for two input EXOR-EXNOR design.

Table VI. Comparison of Individual Gates with Proposed Multi-Logic Gate designs

Gate	No. of Cells with individual gates	No. of cells with multi-logic gates	Approx. consumed area with individual gates (μm^2)	Approx. consumed area with multi-logic gates (μm^2)	Latency of multi-logic gates (sec.)
AND-OR	10	10	13.4	7.6	5
AND-NAND	19	11	23.9	15.7	6
AND-NOR	19	11	23.9	15.7	7
OR-NOR	19	11	25.3	14.7	6
OR-NAND	19	11	25.3	14.7	6
NOR-NAND	28	10	13.4	7.6	5
EXOR-OR	14	12	16.9	11.1	9
EXOR-AND	14	14	15.5	14.4	8
EXNOR-OR	14	14	16.9	14.4	8
EXNOR-AND	14	12	15.5	11.1	9
EXOR-EXNOR	20	16	21.7	26	10

In order to calculate the energy dissipation analysis we employ QCA Designer-E [33-36]. For most of the multi-logic designs, energy dissipation is either comparable or less than the energy dissipation by sum of individual gates as shown in Figure 15. EXOR-OR, EXOR-OR, EXNOR-OR, EXNOR-AND and EXOR-EXNOR multi-logic designs have much reduced power consumption than the combination from individual gates. NOR- NAND design is efficient in terms of power and area consumed both. In EXOR-EXNOR area is compromised but energy dissipation gets reduces.

V. CONCLUSION

In this paper, the proposed concept of efficient multi-logic gates using QCA has been extended. Multi-logic gates are useful in complex digital design. The advantage of multi-logic gates is reduced total QCA cell count compared to sum of individual logic gates. Another benefit of multi-logic gates is that it consumes less area due to its compact design. Power analysis of these multi-logic gates shows that they consume comparable power as individual logic gates. Yet in this study we demonstrate only few multi-logic gates with only 2 inputs and 2 outputs to present the idea of multi-logic-gates using QCA, but this can be extended for more than 2 input and multiple output gates. This work laid the foundation of designing more area and energy efficient QCA designs.

ACKNOWLEDGEMENT

Authors' thanks Department of Electrical and Electronics Engineering at BITS Pilani Dubai Campus for computational resources and other facilities provided.

REFERENCES

[1] G. E. Moore, "Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff.," in IEEE Solid-State Circuits Society Newsletter, vol. 11, no. 3, pp. 33-35, Sept. 2006.

[2] Datta, Suman. "Recent Advances in High Performance CMOS Transistors: From Planar to Non-Planar." (2013).

[3] Lorenzo, Rohit and Saurabh Chaudhury. "Dynamic Threshold Sleep Transistor Technique for High Speed and Low Leakage in CMOS Circuits." Circuits, Systems, and Signal Processing 36 (2017): 2654- 2671

[4] Hussain, A. Singh and S. Chaudhury, "A Review on the Effects of Technology on CMOS and CPL Logic Style on Performance, Speed and Power Dissipation," 2018 IEEE Electron Devices Kolkata Conference (EDKCON), Kolkata, India, 2018, pp. 332-336

[5] Hussain, Inamul & Chaudhury, Saurabh. (2020). A comparative study on the effects of technology nodes and logic styles for low power high speed VLSI applications. International Journal of Nanoparticles. 12. 122. 10.1504/IJNP.2020.10027806.

[6] Semiconductor Industry Association, "International Technology Roadmap of Semiconductors, 2008 Update". [Online]. Available:http://public.itrs.net, 2008

[7] C. S. Lent and P. D. Tougaw, "A device architecture for computing with quantum dots,"in Proceedings of the IEEE, vol. 85, no. 4, pp. 541-557, April 1997.

[8] P.D. Tougaw, C.S. Lent, Logical devices implemented using quantum cellular automata, J. Appl. Phys. 75 (1994) 1818–1825,http://dx.doi.org/10.1063/1.356375.

[9] Porod W, Lent CS, Bernstein GH, Quantum cellular automata. In: Annual report for the period 1 July 1993–30 June 1994, Department of Electrical Engineering University of Notre Dame

[10] M. Graziano, M. Vacca, A. Chiolerio and M. Zamboni, "An NCL- HDL Snake-Clock-Based Magnetic QCA Architecture," in IEEE Transactions on Nanotechnology, vol. 10, no. 5, pp. 1141-1149, Sept. 2011.

[11] Esam AlKaldy, Ali H. Majeed, Mohd Shamian bin Zainal, Danial Bin MD Nor, "Optimum multiplexer design in quantum-dot cellular automata", Indonesian Journal of Electrical Engineering and Computer Science, pp. 148-155, January 2020.

[12] V. Nath, P. K. Barhai, and D. K. Verma, "QCA and CMOS Nanotechnology Based Design and Development of Nanoelectronic Security Devices with Encryption Schemes," TELKOMNIKA Indonesian Journal of Electrical Engineering, pp. 270-279, 2015.

[13] Ali H. Majeed, Esam AlKaldy, Mohd Shamian bin Zainal, Danial Bin MD Nor, "A new 5-input majority gate without adjacent inputs crosstalk effect in QCA technology", Indonesian Journal of Electrical Engineering and Computer Science, pp. 1159-1164, June 2019.

[14] J. C. Das, D. De, S. P. Mondal, A. Ahmadian, F. Ghaemi and N. Senu, "QCA Based Error Detection Circuit for Nano Communication Network," in IEEE Access, vol. 7, pp. 67355-

67366, 2019.

[15] E. Blair, "Electric-Field Inputs for Molecular Quantum-Dot Cellular Automata Circuits," in IEEE Transactions on Nanotechnology, vol. 18, pp. 453-460, 2019. doi: 10.1109/TNANO.2019.2910823.

[16] Lent, C. S., & Tougaw, P. D. (1993). Lines of interacting quantum dot cells: A binary wire. Journal of Applied Physics, 74(10), 6227- 6233

[17] J. Henry, J. Previti, and E. P. Blair, "Electric-field bit write-in for clocked molecular quantum-dot cellular automata circuits," in Proc. 3rd IEEE Int. Conf. Rebooting Comput., Tysons, VA, USA, Nov. 2018, pp. 1–9.

[18] "Implementation of Ex-OR gate using QCA with NNI logic," 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, 2017, pp. 2027- 2031.

[19] Lee Ai Lim, A. Ghazali, S. C. T. Yan and Chau Chien Fat, "Sequential circuit design using Quantum-dot Cellular Automata (QCA)," 2012 IEEE International Conference on Circuits and Systems (ICCAS), Kuala Lumpur, 2012, pp. 162-167.doi: 10.1109/ICCircuitsAndSystems.2012.6408320.

[20] S. K. Anumula and X. Xiong, "Design and Simulation of 4-bit QCA BCD Full-adder," 2019 IEEE Long Island Systems, Applications and Technology Conference (LISAT), Farmingdale, NY, USA, 2019, pp. 1-6.doi: 10.1109/LISAT.2019.8817331

[21] V. Dhare and U. Mehta, "A Simple Synthesis Process for Combinational QCA Circuits: QSynthesizer," 2019 32nd International Conference on VLSI Design and 2019 18th International Conference on Embedded Systems (VLSID), Delhi, NCR, India, 2019, pp. 498-499.doi: 10.1109/VLSID.2019.00105.

[22] Michael Thaddeus Niemier, B.S., "DESIGNING DIGITAL SYSTEMS IN QUANTUM CELLULAR AUTOMATA", A Thesis Submitted to the Graduate School of the University of Notre Dame.

[23] B. Isaksen and C. S. Lent, "Molecular quantum-dot cellular automata," 2003 Third IEEE Conference on Nanotechnology, 2003. IEEE-NANO 2003., San Francisco, CA, USA, 2003, pp. 5-8 vol.2.doi: 10.1109/NANO.2003.1231700

[24] Bahar N, Waheed S, Hossain N, Saduzzaman M. A novel 3-input XOR function implementation in quantum dot-cellular automata with energy dissipation analysis. Alexandria Engineering Journal. 2017;56:1-9. DOI: 10.1016/j.aej.2017.01.022

[25] Avinashkumar,Anuj Borkute, and Nilesh Goel, "An Efficient Design of Multi-logic Gates Using Quantum Cellular Automata Architecture" Springer Nature Singapore Pte Ltd. 2020 ,Modelling, Simulation and Intelligent Computing, Lecture Notes in Electrical Engineering 659, https://doi.org/10.1007/978-981-15-4775-1_67

[26] M. Raj and L. Gopalakrishnan, "High Speed Memory Cell with Data Integrity in QCA," 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2019, pp. 926-929.

[27] Bisma Bilal, Suhaib Ahmed, Vipan Kakkar,"An Insight into Beyond CMOS Next Generation Computing using Quantum-dot Cellular Automata Nanotechnology", International Journal of Engineering and Manufacturing(IJEM), Vol.8, No.1, pp.25-37, 2018.DOI: 10.5815/ijem.2018.01.03

[28] "Radhouane Laajimi, "Nanoarchitecture of Quantum-Dot Cellular Automata (QCA) Using Small Area for Digital Circuits", intechopen, 2018.

[29] Sarkar, Tamal. (2013). Design of AND-OR Gate using Quantum Cellular Automata and its Cell-Cell Response using Computer Simulation.

[30] R. Sen et al., "Priority encoder using reversible logic gates in QCA," 2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, 2017, pp. 319-323.

[31] B. Bilal, S. Ahmed and V. Kakkar, "Multifunction reversibile logic gate: Logic synthesis and design implementation in QCA," 2017 International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, 2017, pp. 1385-1390.

[32] K. Walus, T. J. Dysart, G. A. Jullien and R. A. Budiman, "QCADesigner: a rapid design and Simulation tool for quantum-dot cellular automata," in IEEE Transactions on Nanotechnology, vol. 3, no. 1, pp. 26-31, March 2004.

[33] J. Timler, C.S. Lent, Power gain and dissipation in quantum-dot cellular automata, J. Appl. Phys. 91 (2002) 823–831, http://dx. doi.org/10.1063/1.1421217.

[34] F. Sill Torres, R. Wille, P. Niemann, and R. Drechsler, "An energy- aware model for the logic synthesis of quantum-dot cellular automata," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 37, no. 12, pp. 3031-3041, Dec. 2018.

[35] S. Srivastava, S. Sarkar, S. Bhanja, Estimation of upper bound of power dissipation in QCA circuits, IEEE Trans. Nanotechnol. 8 (2009) 116–127, http://dx.doi.org/10.1109/ TNANO.2008.2005408.

[36] S. Srivastava, A. Asthana, S. Bhanja, S. Sarkar, QCAPro – an error- power estimation tool for QCA circuit design, in: 2011 IEEE International Symposium of Circuits and Systems (ISCAS), 2011, pp. 2377–2380, http://dx.doi.org/10.1109/ ISCAS.2011.5938081

Tungsten as an Interconnect Material for Next-Generation IC Design

Krishna Ganesh
Department of Electrical and Electronics Engineering
Birla Institute of Technology and Science Pilani
Dubai Campus, UAE
E-mail: KrishnaG5713@gmail.com

Vilas H Gaidhane
Department of Electrical and Electronics Engineering and APPCAIR
Birla Institute of Technology and Science Pilani
Dubai Campus, UAE
E-mail:vilasgd612@gmail.com

***Abstract*—Over the last few decades, technology have seamlessly transfer from Aluminium to Copper, as aluminium has deemed as an unsuitable metal when it is in the nanoscale due to associated drawbacks. Although with the challenge of having to replace copper, it has been one of the most consistent materials in the wire dimension area. However, getting a material as ideal as copper is proving to be more inefficient as scale down the wire deeper. There has been a spurt in the research of Tungsten (W) and its Oxides (WOx) as a potential replacement to the issues posed by copper. In general, tungsten has been proven to be superior to most existing interconnect materials and has been only put behind Silver and Copper. However, as technology reach the nanoscale, there has been increasing studies on how to fabricate tungsten and its subsidiaries as an alternate to other materials. The purpose of this research is to study Tungsten and other existing interconnect materials and compare them on theoretical and observational basis. The number of simulations has been carried out using the COMSOL software in various environments in order to see how tungsten behaves in all environments and propose whether it is suitable in the upcoming generation of ICs and beyond.**

Keywords—Nanoelectronics, Interconnects, Tungsten, Conductivity, Resistivity, Integrated Circuits, Conducting materials.

I. INTRODUCTION

There has been a considerable growth over the years in the research of nanowires and other related structures. The reach of this field has touched the disciplines of solid-state chemistry, quantum mechanics, biomedical engineering, condensed matter physics, photonics, opto-electronics etc. However, one of the fields that have benefitted the most is the Nano-electronics/Nanomachines and Integrated Circuit (IC) design. As the IC scale down, the real challenge lies in reducing the dimension of the wire itself. Since the onset of the industrial revolution and even in instances before, wires have been a part of our daily lives for centuries together. Powerlines are utilized efficiently to transmit electricity to households everywhere, the optical fibres have been introduced to transmit information over long and short distance. It has been more than a decade that the researchers all over the world have investigated different processes and techniques and have come up with methods which can be used to synthesize nanowires which are thousands of times thinner than the hair. These nanowires have allowed the onset of the next generation of application in fields like advanced computing, photonics and photosensitive applications, efficient energy utilization and especially it has enhanced applications in the biomedical field as well [1-5].

The material which have been scaled down to the nanoscale in the form of nanotubes, nanorods and nanowires have been under the spotlight due to their plethora of applications in semiconductors, interconnects and other nanoscale devices. These materials give the insight on the study of properties like electrical, mechanical and thermal which have changed due to scaling down to the nanoscale. These properties can be studied and altered for development and advancement of novel applications in different fields and focusing on the nanoelectronics.

Choi and Barmak [6], in their research have discussed the potential of Tungsten as future interconnects and have reviewed the current availability of different technology, processing methods and the required nanostructure. Moreover, they have studied the resistivity Copper and Tungsten nanowires using the Mayadas-Shatzkes grain boundary and Fuchs-Sondheimer surface scattering model model. It is observed that Tungsten nanowires exhibit lower resistivity than Copper. Bien *et al.* [7] have extensively studied a synthesis method for fabricating Tungsten nanowires which are self-aligned, by the use of a Poly-Si core. They formed nanowires in the neighbourhood of 10nm using the Poly-Si transfer technology followed by CVD of tungsten using WF6 as the precursor solution. They also characterized the conductivity of the nanowires to see how this method influenced physical properties and found it to be 40% higher than doped Poly-Si nanowires of the same scale. Further, Wu *et al.* [8] has proposed tungsten oxide materials as a photothermal nanomaterial. They found that WO_x structures have an unusual oxygen defect and allow excellent photo-absorption of various wavelengths in the NIR region. Moreover, Min and Ahn [9,10] have reported the synthesis of W nanostructures by a simple process of thermally treating tungsten films, which are characterized by self-catalysing layer formation and have astounding field emission (FE) properties. This easy fabrication could be the gateway for building nano-level interconnection materials and nano-machine components. Chen *et al.* [11] have come up with a novel method which involves the usage of solution-based printing of thermoelectric generators (TEG) and metal contacts and Ni *et al.* have also discussed the improvement in thermoelectric properties using a PEDOT nanowire film to coat Te, which can also be investigated in the use case of tungsten [12].

From above literature, it is observed that the Tungsten has been extensively investigated and utilized for various applications as an interconnect. Moreover, it can be used in energy applications, biomedical, chemical, photothermal etc. The versatility of tungsten, along with its inter-miscibility

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

with other metals to form alloys can help to investigate oxides of W, Carbides and even Sulphides as an alternative. Therefore, in this paper an extensive experimentation is carried out the possibilities of Tungsten as an interconnect for IC design has been investigated.

The rest of the paper is structured as follows: Section II describes the detailed explanation of the existing material and characteristics. Experimental evaluation and results are presented in Section III and conclusions are drawn in Section VI.

II. MATERIALS AND THEIR CHARACTERISTICS

This sub-heading explores the current available information on the current generation methods and the viable synthesis methods for such nanowires. There are a variety of metals available, which can be used for IC and Semiconductor design. Some metals which are commonly used are Silicon, Copper, Aluminium, Silver and Gold.

Silicon (Si) is the most commonly used material for IC design. Si is a semiconductor having the electrical behaviour is between that of a conductor and an insulator at room temperature. Polycrystalline silicon (Poly-Si) is used for making resistors or conductors in an integrated circuit. Si is also widely used in semiconductor packaging, being the main ingredient of plastic encapsulants for ICs and are used in die overcoats as well [13,14].

Another material Aluminium (Al) is a highly lightweight metal with silvery appearance. It is the most abundant metallic element on earth and used in many aspects of semiconductor manufacturing. On an IC, Al metal lines are commonly used as the main conductor between components due to low resistivity. As a thin film, it has good adherence to SiO_2. Al is used for the bonding and probing pads on the die. When used for IC metallization, it slightly doped with elements such as Si and/or Cu to improve its characteristics and reliability. In semiconductor assembly, ceramic packages are composed mainly of alumina. Al is also used for wire-bonding IC in ceramic packages. Aluminium is one of the most widely used metal for metallizing IC chips since it has a high conductivity, very low contact resistance and good compatibility to Si and SiO_2. However, Aluminium has many shortfalls when it is brought up in the interconnect argument like poor step coverage, spiking and migration of electrons [13-16].

Gold (Au), is a soft element which is a good conductor of heat and electricity. It is also the most malleable and ductile of all metals. Au is used in many aspects of semiconductor manufacturing, particularly in the assembly or packaging processes. Its most widespread use is in wire-bonding because of excellent conductivity and ductility. It is mainly used as wires for interconnect in ICs [17].

Silver (Ag), is another element which is next only to gold in terms of malleability and ductility and is also a good conductor of heat and electricity. Ag is the best conductor of electricity, better even than copper and gold. Ag is used in many facets of semiconductor manufacturing, more particularly in packaging processes. Most epoxy die attach materials contain Ag fillers for increased electrical and thermal conductivity. Ag is also used to cover the surfaces of the die pad and bonding fingers of the plastic packages to prevent chemical degradation of these areas, which may lead to die attach and bonding problems [18].

Copper (Cu), is one of the most widely used metals in the world, mainly because of its many desirable properties. It is the 2nd best conductor of electricity next to Ag. It is very malleable and ductile and good conductor of heat. Cu is also widely used in semiconductor assembly being an excellent conductor. Recent technological advancements though have already allowed the use of copper as metal lines in semiconductor devices [19]. The continued scaling down of semiconductor devices is proportional to the miniaturization of Cu interconnect lines as well. Although this scaling leads to improved working, it tends to exhibit a higher resistivity and hence in turn give rise to issues like increased power consumption, delayed signal transfer and increased resistivity.

With the onset of advance techniques of VLSI on circuit chips, the introduction of new processes and materials which are capable of matching or improving the density of packing and high-performance are remains the challenge for researchers. There is a need of a material which can help in shunting out Poly-Si gate and thereby reduce the sheet resistance, electro-migration, junction spiking and a high contact resistance.

From literature, it has been observed that the Tungsten can be considered as a potential candidate to replace Cu for metallization for two main reasons. The first is the very high melting point (3695) compared with Cu, which is expected to improve interconnect reliability such as electromigration and stress. The second reason is the anticipated reduction of resistivity effect due to significantly shorter EMFP of 19.1 nm at 293˚K as compared to Cu. However, given the fact that the room-temperature bulk resistivity of W at 5.3 μΩ-cm is higher than that for Cu at 1.7 μΩ-cm, it is important to quantify the contributions of size-dependent scattering mechanisms in order to determine whether W can favourably compare with Cu as a nanoscale interconnect. The theoretical nanowires properties of existing interconnect material is summarised in Table 1.

The metallic Tungsten has well-known chemical, physical, electrical, and mechanical properties that makes it very useful for various applications [20-22]. Especially for nanodevices below 100 nm, a gate material such as W with a mid-gap work function is desirable as a metal gate of CMOS technology due to the fairy low resistivity. Another variant of tungsten such as its subsidiary oxides are used commonly in semiconductor and nanotechnology-based applications. The research potential in WO_x can be seen as early as the 17th century, when people began studying the different properties of $LiWO_3$ and were searching for methods to grow compounds like WO_3 and $NaWO_3$. Along with the progress of nanotechnology, the fabrication of Tungsten and its oxide have become increasingly prominent as nanostructures of W and WO_x have better performance. Nano-structured WO_x is exceptionally versatile and offers a wide range of unique characteristics when it comes to VLSI design. In comparison to other oxides such as TiO_2, ZnO, NiO, WO_x exhibits much more advanced electrical properties and also being employed as material which can be used in nano-sensing [23-25]. Thus, it has been observed that the Tungsten and its oxides can be used for interconnect in integrated circuits design.

In this paper, an attempt is made and the various existing materials as well as Tungsten oxides in the form of nanowires has been simulated using the COMSOL software. Moreover,

the detailed characteristics of simulated nanowires has been studied.

TABLE I. THEORETICAL NANOWIRE PROPERTIES OF EXISTING INTERCONNECT MATERIALS

Parameters	Aluminium (Al)	Copper (Cu)	Gold (Au)	Silver (Ag)	Silicon (Si)
Molecular Wt.	26.98	63.55	196.97	107.87	28.086
Melting Point (^{0}C)	660.37	1085	1064	961.8	1414
Boiling Point (^{0}C)	2467	2562	2700	2162	3265
Density (g/cm^3)	2.7	8.96	19.3	10.49	2.33
Resistivity	2.6548 μΩ-cm	1.673 μΩ-cm	24.4 nΩ-cm	15.9 nΩ-cm	64 mΩ-cm
Electronegativity	1.5	1.90	2.54	1.93	1.9
Heat of Fusion (kJ/mol)	10.79	13.26	12.55	11.3	50.55
Heat of Vaporization (kJ/mol)	293.4	300.4	334.4	250.58	384.22
Poisson's Ratio	0.35	0.34	0.42-0.44	0.337	0.22
Specific Heat (kJ/kg K)	0.900	0.390	0.129	0.240	0.710
Thermal Conductivity (Wm^{-1} K)	205	401	310	406	150
Coefficient of Thermal Expansion (μm-m^{-1} K^{-1})	23.1	16.5	14	18	2.6
Vicker Hardness (MPa)	167	369	216	251	9630
Young's Modulus (GPa)	0	110-128	79	85	130-180
Tensile Strength (MPa)	40-700	210	80-700	100-200	20-170

III. EXPERIMENTATIONS AND OBSERVATIONS

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar. In this experiment, the various experimentation is carried out using COMSOL Multiphysics Modelling on the system which has Intel(R) Core (TM) i7 CPU with 2.4 GHZ frequency processor and 8 GB RAM.

In the first experiment, the Aluminium, Coper and Silver wire has been simulated with the various parameters. Here, the "Electric Currents" module on COMSOL version 5.3a has been used to simulate the particular wire. The available quantities such as resistance, reference impedance, terminal voltage, terminal current, total electrical energy as shown in Table 2. Moreover, the current of $1\mu A$ through the wire with dimension radius 0.5nm, and length 10nm is considered for the evaluation. Table 3 shows the calculated values of resistivity, conductivity and current density, respectively. The details of simulation of Aluminium, copper and silver using COMSOL software is shown in Fig 1.

TABLE II. OPERATIONAL VALUES FOR ALUMINIUM, COPPER AND SILVER

Parameters	Aluminium	Copper	Silver
Resistance (Ω)	370	233	226.52
Electrical Energy (J)	4.337×10^{-29}	1.717×10^{-29}	-2.48×10^{-28}
Reference Impedance (Ω)	50	50	50
Terminal Voltage (V)	3.69×10^{-4}	2.326×10^{-4}	2.26×10^{-4}

TABLE III. CALCULATED VALUES OF RESISTIVITY, CONDUCTIVITY AND CURRENT DENSITY

Parameters	Al	Copper	Silver
Resistivity (Ω-m) $\rho = RA/l$	2.91×10^{-8}	1.83×10^{-8}	1.78×10^{-8}
Conductivity (S/m) $\sigma = 1/\rho$	3.4×10^{7}	5.46×10^{7}	5.6×10^{7}
Current density (A/m^2) $J = I/A$	3.82×10^{12}	3.82×10^{12}	3.82×10^{12}

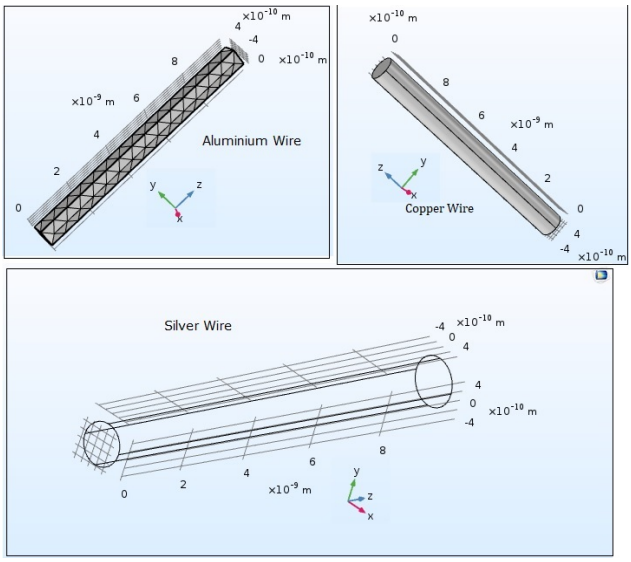


Fig. 1. Simulation of Aluminium, Copper and Silver wire.

In second experiment, the normal tungsten with same dimensions as previous experiment has been simulated and various parameters are obtained. The resistivity ($6.1\text{x}10^{-8}$

Ωm), conductivity (1.64x10^7 S/m), and current density (3.82x10^{12}A/m^2) which is better than the other materials such as Aluminium, copper and silver. The simulation of Tungsten nanowire is shown in Fig. 2.

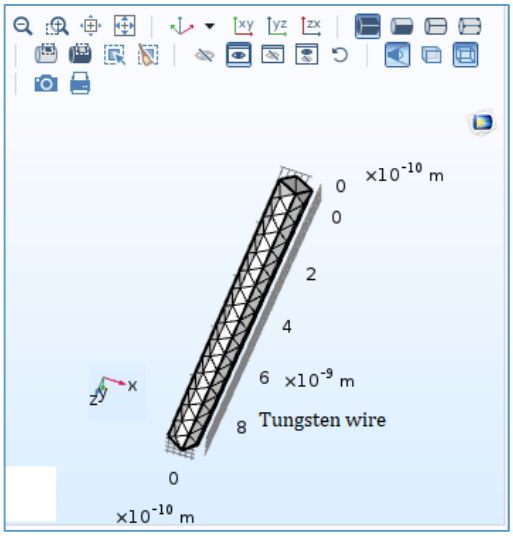


Fig. 2. Simulation of Tungsten Nanowire

From the study of Aluminium, copper, silver and Tungsten, it cannot be concluded whether Tungsten material is viable as an interconnect. Therefore, it has been decided to study the different types of Tungsten Alloy in COMSOL to find the suitable material nanowire interconnect. Following materials are considered and simulated in COMSOL:

- Tungsten [CVD <001>, tested at 293K, 9500/s]
- Tungsten [CVD <011>, tested at 293K, 6600/s]
- Tungsten [gas]
- Tungsten [liquid]
- Tungsten [MEMS]
- Tungsten [powder metallurgy, tested at 293K, 4000/s, annealed at 2873K]
- Tungsten [rod, tested at 3073K]
- Tungsten [sheet,1.5 mm sheet, strain relieved at 1423K]
- Low Carbon Steel Tungsten Steel
- Tungsten [swaged bar, tested at 1033K]
- Tungsten [Built in]

There are mainly two parameters in the COMSOL material contents toolbar, without which the simulation will not run properly. The first is relative permittivity of 6.2438 and 22.254 and second is electrical conductivity 1.79×10^7 *S/m*. These both values have been obtained from online databases and verified. We have simulated all of tungsten available in the COMSOL material library and a separate file is prepared for each material to ensure no discrepancies in the observed values. Following values of parameters has been used in each calculation:

- Nanowire Radius: 0.5 *nm* or 0.5×10^{-9} *m*
- Nanowire height: 10 *nm* or 10×10^{-9}*m*
- Current: 1*nA*
- Electrical conductivity: 1.79×10^7 or 1.79×10^7 *S/m*
- Relative permittivity: ε_1= 6.2438, ε_2 = 22.254.

The resistivity of the materials has been calculated using the formula:

$$\rho = \frac{RA}{L} \tag{1}$$

where, *R* is resistance, *A* area of cylindrical wire and *L* is the length per height of the wire. The various simulations have been carried out using COMSOL for different Tungsten variant. The simulation has been shown in Fig. 3. The simulation results are summarized in Table 4. The simulated tungsten wires for a nanowire yielded a resistivity of 6.8×10^{-8} in most of the cases. However, this is insufficient information as to whether top-down manufacture will influence the resistivity of the metal. Since we had to provide the values of conductivity and the relative permittivity values in each case, we found that the former had a bigger impact on the resistance value in each case rather than the latter. The tungsten liquid has a higher resistance in liquid which is attributed to the fact that tungsten is covalent in nature and considering the fact that ionic solutions have higher conductivity. The Tungsten [MEMS] structure shows a slightly lower resistance as compared to the built-in Tungsten. This value also corresponds to the theoretical obtained resistivity value.

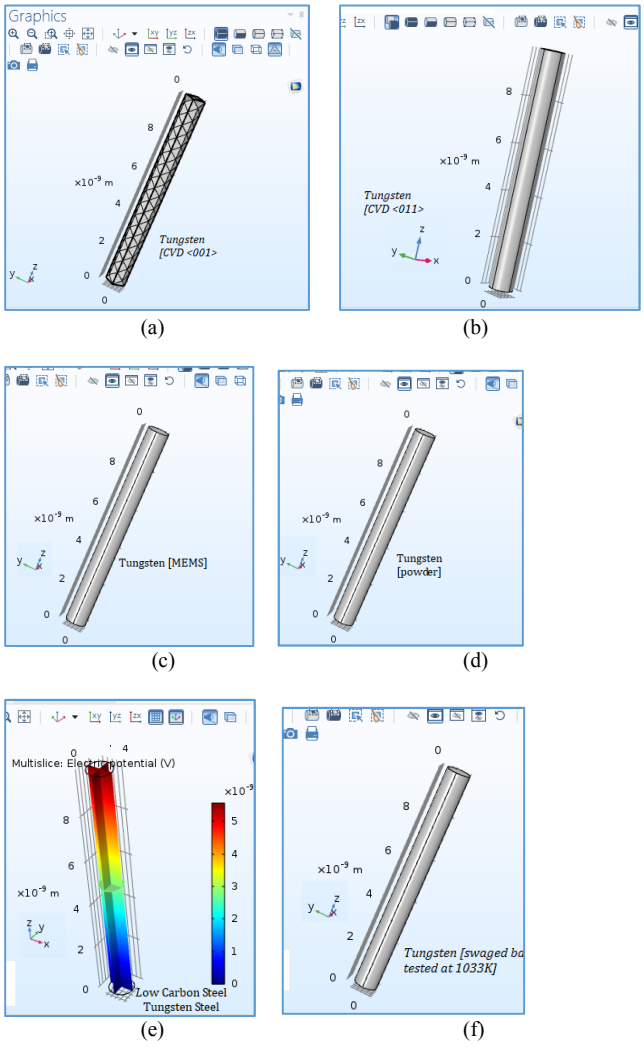


Fig. 3. Simulation of Tungsten Nanowire; (a) Tungsten [CVD <001>], (b) Tungsten [CVD <011>], (c) Tungsten [MEMS], (d) Tungsten [Powder], (e) Low Carben Steel Tungsten Sheet, (f) Tungsten (Swaged).

TABLE IV: CALCULATED VALUES OF RESISTIVITY, CONDUCTIVITY AND CURRENT DENSITY

Material	Electrical Parameters					
	Resistance (Ω)	Total Energy (J)	Reference Impedance (Ω)	Terminal Voltage (V)	Resistivity (Ω)	Conductivity (S/m)
Aluminium	370	4.3×10^{-29}	50	3.6×10^{-4}	2.91×10^{-8}	3.4×10^{7}
Copper	233	1.7×10^{-29}	50	2.3×10^{-4}	1.83×10^{-8}	5.4×10^{7}
Silver	226.52	-2.48×10^{-28}	50	2.2×10^{-4}	1.78×10^{-8}	5.7×10^{7}
Tungsten	775	1.91×10^{-28}	50	7.7×10^{-7}	6.1×10^{-8}	1.6×10^{7}
Tungsten (CVD 001)	780	4.3×10^{-33}	50	7.8×10^{-7}	6.123×10^{-8}	1.63×10^{7}
Tungsten (CVD 011)	779	1.2×10^{-33}	50	7.7×10^{-7}	6.1×10^{-8}	1.6×10^{7}
Tungsten (Gas)	779	4.29×10^{-33}	50	7.8×10^{-7}	6.1×10^{-8}	1.6×10^{7}
Tungsten (Liquid)	18209	2.34×10^{-30}	50	1.8×10^{-5}	1.43×10^{-6}	7×10^{5}
Tungsten (MEMS)	690	3.4×10^{-33}	50	6.97×10^{-7}	5.4×10^{-8}	1.85×10^{7}
Tungsten (Powder)	779	4.29×10^{-33}	50	7.8×10^{-7}	6.1×10^{-8}	1.6×10^{7}
Tungsten (Rod)	779	4.29×10^{-33}	50	7.8×10^{7}	6.1×10^{-8}	1.6×10^{7}
Tungsten (Sheet)	779	4.29×10^{-33}	50	7.8×10^{-7}	6.1×10^{-8}	1.6×10^{7}
Tungsten (Steel Alloy)	5.6	9.8×10^{-33}	50	5.58×10^{-9}	4.4×10^{-10}	2.3×10^{9}
Tungsten (Swaged Bar)	779	4.29×10^{-33}	50	7.8×10^{-7}	6.1×10^{-8}	1.6×10^{7}

The Tungsten steel alloy had different results. Typically, alloys have to exhibit physical, chemical and electrical constants in relation to their parent metal. It has to be in between both the metals or closer to the base metal as per principle. The experiment started with 2.5×10^{6} S/m as the first value of electrical conductivity in order to obtain resistance and other parameters. The obtained resistance value is 780Ω, which is similar to the resistance of tungsten. When the conductivity is increased to 10^{7}-10^{8}, the resistance remains the same, which indirectly implies that Tungsten is the base metal. However, once we used 2.3×10^{9}, the resistance dropped to approximately 6Ω. Hence in order to get a lower resistivity, an alloy should be used to of a metal to reduce the resistivity or increase the conductivity. This intuition can be used to determine an alloy would serve as a better interconnect material. Another observation which plays a large role in design of Interconnects is the down scaling of the wires. The presented experimental data indicates that the resistivity in metal based thin films and wires in the nanometre dimension which is inversely related to the width/diameter. This is mainly attributed to the electron scattering which has increased due to Grain Boundaries (GB) and the roughness of the metal Surface Roughness (SR) of the wire boundaries. In simple words, if size decreases, resistance increases and vice versa. This in turn, has an effect wherein the wire dimensions have the constraint of scaling down along with the IC or Transistor size, and this scaling down of resistivity also gives rise to other shortcomings like higher heating issues, increased power consumption and delay in signal transmission. If this increase in resistivity is not kept in check, the interconnects will serve as the main issue during the fabrication of next generation ICs. Existing experimental data available tells us that the resistivity of these wires in the nanoscale agree with the Fuchs-Sondheimer (FS) and Mayadas-Shatzkes (MS) models and these models specifically utilize the fitting parameters and along with approximation of semi-classical mechanics. However, these models fail to provide the data about the domain of quantum mechanics when it comes to scattering, confinement and a well-defined and mapped structures of SRs and GBs. These models make use of fitting parameters and semi-classical approximations and do not provide further insight in the quantum-mechanical effects of scattering, confinement and the detailed structure of the GBs and SR. From studies, however, there has been a marginal decrease in the resistance. When the transition from the microscale to the nanoscale is considered, the classical mechanics will be replaced by the principles of quantum mechanics. The effect on the resistance and other physical properties need to be studied more in detail as COMSOL provides a limited insight. Since these properties cannot be altered, it is necessary to find a suitable alloy or mixture of tungsten which can exhibit the expected characteristic. It has been observed that tungsten alloys have extremely high corrosion resistance on par with Titanium when it comes to corrosion and stress. Moreover, it is highly durable with higher melting point (3422^{0}C) and boiling point (5555^{0}C).

IV. CONCLUSIONS

In this work, an extensive study has been carried out on the different materials for the nanowire's applications in integrated circuits design. The Tungsten and its variants are considered in this study and their electrical characteristics has been calculated using the various simulations. The obtained results are compared with mostly used materials like Aluminium, Copper, Silver, Silicon and Gold. It has been observed that the Tungsten alloy has better conductivity and resistivity as compared to the existing materials. Thus, the Tungsten alloys can be used as an interconnect in integrated circuits.

As a future scope for thsese studies, it can be further investigated tungsten to come to a more conclusive result. After noticing the properties of tungsten and silver, the research is being focused towards coming up with a tungsten-silver alloy as the next generation interconnect material. The alloy can be formed by either mixing both the molten metals and allowing alloy formation or powder metallurgy which

involves the sintering of powders of both metals. Theoretically, the resistivity of the metal should be close to the base metal, or in between both the metals, based on the percentage of each metal used in the alloy. The alloy may be having a proportion of 35-70% of each metal so as to not lose the base properties of the other. Tungsten and Silver together can overcome each one of the other's drawbacks. Although getting a conductivity close to the value of 2.3×10^9 as seen in Table IV is nearly impossible considering the current technology, working on an alloy which can provide a conductivity closer to Ag is very much possible. It can increase the inertness, melting point, boiling point and even the corrosion resistance of Silver, whereas Ag in turn can increase the conductivity, malleability and ductility of tungsten.

REFERENCE

[1] W. Lu, P. Xie and C. M. Lieber, "Nanowire Transistor Performance Limits and Applications," IEEE Trans. Electron Devices, vol. 55, No. 11, pp. 2859-2876, 2008.

[2] Z. Gu, H. Li, W. Yang, Y. Xia and J. Yao, "Large-scale synthesis of single-crystal hexagonal tungsten trioxide nanowires and electrochemical lithium intercalation into the nanocrystals," J. Solid State Chem., vol. 180, no. 1, pp. 98-105, 2007.

[3] S. J. Yoo, J. W. Lim, Y. E. Sung, Y. H. Jung, H. G. Choi and D. K. Kim, "Fast switchable electrochromic properties of tungsten oxide nanowire bundles," Appl. Phys. Lett., vol. 90, no. 17, pp. 173126, 2007.

[4] J. I. Pascual, J. Méndez, J. Gómez-Herrero, A. M. Baro, N. Garcia, Uzi Landman, W. D. Luedtke, E. N. Bogachek, and H-P. Cheng, "Properties of metallic nanowires: from conductance quantization to localization," Science. vol. 267, no. 5205, pp. 1793–1795, 1995.

[5] P. A. Smith, C. D. Nordquist, T. N. Jackson, T. S. Mayer, B. R. Martin, J. Mbindyo and T. E. Mallouk, "Electric-field assisted assembly and alignment of metallic nanowires," Appl. Phys. Lett., vol. 77, pp. 1399–1401, 2000.

[6] D. Choi and K. Barmak, "On the potential of Tungsten as Next-generation Semiconductor Interconnect," Electron. Mater. Lett., vol. 13, no. 5, pp. 449-456, 2017.

[7] D Bien, R. M. Saman, S. A. M. Badaruddin and H. W. Lee, "Selective formation of tungsten nanowires," Nanoscale Res. Lett., vol. 6. no. 1, pp. 1–5, 2011.

[8] Z. W. Wu, M. Wang, C. W. Guo, Z. W. Shan, J. Li, J. and W. Z. Han, "Graphene-coated tungsten nanowires deliver unprecedent modulus and strength," Mater. Res. Lett., vol. 7, no. 2, pp. 47-52, 2019.

[9] J. D. Meindl, R. Venkatesan, J. A. Davis, J. Joyner, A. Naeemi, P. Zarkesh-Ha, M. Bakir, T. Mule, P. A. Kohl, and K. P. Martin, "Interconnecting device opportunities for gigascale integration (GSI)," In IEEE International Electron Devices Meeting. Technical Digest (Cat. No. 01CH37224), 2001, pp. 1-4.

[10] J. D. Meindl, J. A. Davis, P. Zarkesh-Ha, C. S. Patel, K. P. Martin, and P. A. Kohl, "Interconnect opportunities for gigascale integration," IBM J. Res. Develop., vol. 46, no. 2.3, pp. 245–263, Mar. 2002.

[11] Chen, B., Kruse, M., Xu, B., Tutika, R., Zheng, W., Bartlett, M. D., ... & Claussen, J. C.. Flexible thermoelectric generators with inkjet-printed bismuth telluride nanowires and liquid metal contacts. Nanoscale, 11(12), pp. 5222-5230. 2019

[12] Ni, D., Song, H., Chen, Y., & Cai, K. Significantly enhanced thermoelectric performance of flexible PEDOT nanowire film via coating Te nanostructures. Journal of Materiomics, 6(2), pp. 364-370. 2020

[13] M. Z. Atashbar, S. Singamaneni, "Room temperature gas sensor based on metallic nanowires," Sens. Actuaters B-Chem., vol. 111, pp. 13–21, 2005.

[14] R. Ko, S. Wang, Z. Wen, J. Lin, G. Fan, W. Shu and B. Liou, "Development of gas sensors based on tungsten oxide nanowires in Metal/SiO2/Metal structure and their sensing response to NO2" Jpn. J. Appl. Phys., vol. 72, pp. 32-47, 2008.

[15] C. S. Rout, M. Hedge, C. N. R. Rao, "H2S sensors based on tungsten oxide nanostructures", Sens. Actuator B Chem., vol.12, pp. 2-8, 2008.

[16] P. M. Parthangal, M. R. Zachariah, R. E. Cavicchi, C. B. Montgomery, S. Turner, "Restructuring tungsten thin films into nanowires and hollow square cross-section microducts," J. Mat. Res., vol. 20, no. 11, pp. 11-20, 2011.

[17] S. Chungpaiboonpatana and F. G. Shi, "Packaging of copper/low-k IC devices: A novel direct fine pitch gold wirebond ball interconnects onto copper/low-k terminal pads," IEEE Trans. Adv. Packag., vol. 27, no. 3, pp. 476-489, 2004.

[18] G. Hu, "Comparison of copper, silver and gold wire bonding on interconnect metallization," In 13th IEEE International Conference on Electronic Packaging Technology & High-Density Packaging, 2012, pp. 529-533.

[19] Y. Morand, "Copper metallization for advanced IC: requirements and technological solutions," Microelectron. Eng., vol. 50, no. 1-4, pp. 391-401.

[20] H. Zheng, J. Z. Ou, M. S. Strano, R. B. Kaner, A. Mitchell and K. Kalantar□zadeh, "Nanostructured tungsten oxide–properties, synthesis, and applications," Adv. Funct. Mater., vol. 21, no. 12, pp. 2175-2196, 2011.

[21] M. Rieth, S. L. Dudarev, S. G. De Vicente, J. Aktaa, T. Ahlgren, S. Antusch, D. E. J. Armstrong, M. Balden, N. Baluc, M. F. Barthe and W. W. Basuki, "Recent progress in research on tungsten materials for nuclear fusion applications in Europe," J. Nucl. Mater., vol. 432, no. 1-3, pp. 482-500, 2013.

[22] S. Cong, Y. Tian, Q. Li, Z. Zhao and F. Geng, "Single-crystalline tungsten oxide quantum dots for fast pseudocapacitor and electrochromic applications," *Adv. Mater.*, vol. 26, no. 25, pp. 4260-4267, 2014.

[23] S. H. Wang, T. C. Chou and C. C. Liu, "Nano-crystalline tungsten oxide NO2 sensor," Sens. Actuaters B-Chem., vol. 94, no. 3, pp. 343-351, 2003.

[24] A. Maity and S. B. Majumder, "NO2 sensing and selectivity characteristics of tungsten oxide thin films," Sens. Actuaters B-Chem., vol. 206, pp. 423-429, 2015.

[25] K. Y. Ko, J. G. Song, Y. Kim, T. Choi, S. Shin, C. W. Lee, K. Lee, J. Koo, H. Lee, J. Kim and T. Lee, "Improvement of gas-sensing performance of large-area tungsten disulfide nanosheets by surface functionalization," ACS nano, vol. 10, no. 10, pp. 9287-9296, 2016.

BIG DATA MANAGEMENT OBLIGATION IN SMART GRID TECHNOLOGIES, THE FUTURE OF TOMORROW

1st Dafer Alali
Department of Computer Science and Engineering,
Oakland University,
Michigan, USA
dalali@oakland.edu

2nd Surah Al Dakhl
Department of Electrical and Computer Engineering,
Oakland University,
Michigan, USA
saldakhl@oakland.edu

3rd Mohamed Zohdy
Department of Electrical and Computer Engineering,
Oakland University,
Michigan, USA
zohdyma@oakland.edu

4th Mohammed Mahmoud
Department of Computer Science and Engineering,
Oakland University,
Michigan, USA
mahmoud2@oakland.edu

***Abstract*— A smart grid is an intellectual electricity grid that enhances the production, distribution and consumption of electricity across the insertion of Information and Communication Technologies on the electricity grid. In quintessence, smart grids bring profound modifications in the information systems that drive them, new data streams coming from the electricity grid, new participants such as regionalized producers of renewable energies, new uses such as electric vehicles and connected homes and new communicating equipment such as smart meters, sensors and remote control points. Altogether, this will result in an overflow of information, which the energy corporations will have to handle. Big data technologies propose proper solutions for utilities, although the decision regarding which big data technology to deploy is critical. In this paper, we present an overview of data management for smart grids, summarize the enhanced value of big data technologies for this type of data, and discuss the technical obligations, tools and key steps to employ big data solutions in the smart grid framework.**

Keywords—component, smart grid, management, Big Data, smart meters, renewable energies, electric vehicles

I. INTRODUCTION

Electricity depletion has improved in practice and in nature while electricity usages are growing positive energy structures, electric mobility, adjustable intensity urban lighting, storage batteries, etc. The electricity production modes are also developing thanks to the improvement of renewable energies and the transformation of the energy mix. The electrical system needs to develop towards better reliability, efficiency and flexibility. Smart grids become a great resolution to these concerns, through presenting Information and Communication Technologies (ICT) into electricity grids and integrating effectively the events of all users (manufacturers and customers) in order to assure a sustainable, safe and cost-effective delivery of electricity. Smart grids guarantee effectual connection and utilization of all means of manufacturing, offering automatic and real-time managing of the electrical networks, permit better measurement of spending, enhance the level of consistency and improve the current services which in turn lead to energy savings and lower costs. The application of smart grids features a prime escalation in the volume of data to be handled due to the mechanism of smart meters and several sensors on the network and the improvement of consumer services, etc. For instance, a smart meter can direct the customer energy usage every 15 minutes, therefore every million meters could create 96 million reads per day instead of one meter reading a month in a conservative grid. In addition to energy managing, smart grids need great data managing to be adept to deal with high velocity, storage capacity and advanced data analytics obligations. Certainly, smart grid data involves complex treatments, due to their nature, sharing and real-time restrictions of certain needs. Big data techniques are appropriate for advanced and effectual data management for this type of use. The key purpose of utilities is the capability to manage high data and to deplete advanced analytics to convert data collected to information, then to knowledge and lastly to actionable procedures. This paper presents an overview of the prospects, concepts and challenges of data managing in smart grids with the importance of big data substructure. Additionally, it illustrates the main principles and resources obligations utilities must test in order to choose the right big data tools given data analytics system.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

II. SMART GRID PLATFORM

Smart grids can be abstracted as a huge cyber-physical systems that supports and substantially improves controllability and responsiveness of highly dispersed sources and assets within electric power systems. Renewable generations will make a progressively significant contribution to electric energy production for the future. Integration of these highly variable, widely distributed sources, will call for new attempts to power system operation and control. The European Union's (EU's) smart grid technology platform encapsulates the advantages of smart grids as follows [1]:

1. Improved facilitation of the connection and operation of generators of all sizes and technologies.
2. Tolerate customers to engage in optimizing the operation of the system.
3. Offer customers with greater information and options for choice of supply.
4. Drastically decrease the environmental impact of the whole electrical supply system.
5. Sustain the existing high levels of system reliability, quality and security.
6. Supply, maintain, and improve the existing services efficiently and foster market integration.

III. INFRASTRUCTURE

The infrastructure of a smart grid scheme is continuously relative to fixed objectives and capabilities. The employment of a smart grid can develop the robustness, self-healing capability, and integrality of the grid [2].

Table 1. Domains in the Smart Grid Conceptual Model by NIST.

Domain	Description
1 Customer	Where electricity is consumed. Sub-domains are homes, commercial and industrial customers. Actors may also generate, store and manage energy use.
2 Markets	Where grid assets are exchanged. Actors are the operator and participants in electricity markets
3 Service provider	Where support services for producers, distributors and customers are performed. Actors are organizations providing services to electrical customers and to utilities.
4 Operations	Where proper operation of the power system is ensured. Actors are the managers of the movement of electricity
5 Bulk generation	Where delivery of electricity to customer starts. Actors are the generators of electricity in bulk quantities and may also store energy for later distribution.
6 Transmission	Where bulk transfer of power from generation to distribution is done. Actors are the carriers of electricity over long distances and may also store and generate electricity.
7 Distribution	Where transmission, customer, consumption metering, distributed generation and distributed storage interconnect. Actors are the distributors of electricity to and from customers.

The National Institute of Standards and Technology (NIST) describes a theoretical model that endorses planning, promotes improvement, certification, and the incorporation of interconnected networks and equipment that will combine the smart grid. The National Institute of Standards and Technology (NIST) describes a theoretical model that endorses planning, promotes improvement, certification, and the incorporation of interconnected networks and equipment that will combine the smart grid. NIST has separated the smart grid into seven domains (with subdomains) that include smart grid actors and applications. Moreover, it categorizes actor devices (such as smart meters and solar energy generators), systems (such as control systems), programs, and stakeholders that make evaluations and exchange information necessary for performing applications; applications as tasks executed by one or more actors within a domain (such as home automation, solar energy generation, and energy storage and energy management). A distribution utility, for example, will include actors in the operations domain, such as a distribution management system, and in the customer domain, such as electric meters.

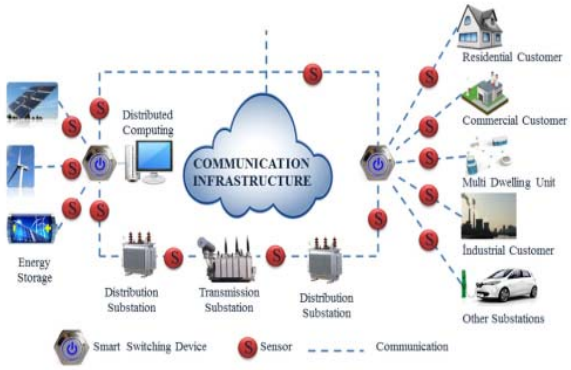


Fig. 1. Smart grid infrastructure [3].

IV. SMART GRID FUNCTIONALITIES

The smart grid implies answers and solutions to the electricity supply adequacy concerns. The Energy Independence and Security Act of 2007 (EISA) lays the stage for the reconstruction of the electricity grid. The smart grid section lists the following characteristics [4]:

A. Reliability, security, and efficiency of the electric grid

A reliable power supply is critical to any power system. It concludes the accomplishment of the grid in delivering the needed service to the end customers. As grids continue to grow in size and complication, it has become more challenging to analyze grid reliability, but new analytical approaches from research attempts have continued to develop a stronger reliability substance for modern networks. A data mining procedure to detect grid system structure from raw historical system data can evaluate grid service reliability by using Bayesian networks [5]. Remote monitoring of hybrid generation and automatic smart grid management for

unsteady distribution may influence efficiency [6]. Smart grid systems risk analysis and integrated Systems Security Threat Model (SSTM) that assist to better understand the weaknesses exploited by attackers [7].

B. Implementation and integration of distributed sources and generation

Distributed energy resources (DER) are small supplies of power that can help encounter regular power demand. DER such as storage and renewable technologies expedite the transition to smart grids [8]. The coming in of renewable energy resources such as distributed generators can help mitigate the difficulties of diminishing fossil reserves and the increasing customer demand. Distributed generations that involve wind generators, photovoltaic generators, and battery storage systems could feature thermal generation and electric vehicles [9]. The combination of these resources, though, also denotes that remarkable amounts of data would require be controlling and processing.

C. Demand response and demand-side sources

The Federal Energy Regulatory Commission (FERC) expresses demand response as "alterations in electric use via demand-side sources from their normal expenditure patterns in response to alterations in the cost of electricity over time, or to incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized" [10]. Demand response offers customers the possibility to be engaged in grid operations as they can decrease or change their electricity usage through ultimate periods and profit via financial incentives. The growth of grid transformation technologies and techniques for demand response is one of the objectives of the US Department of Energy (DOE) [11].

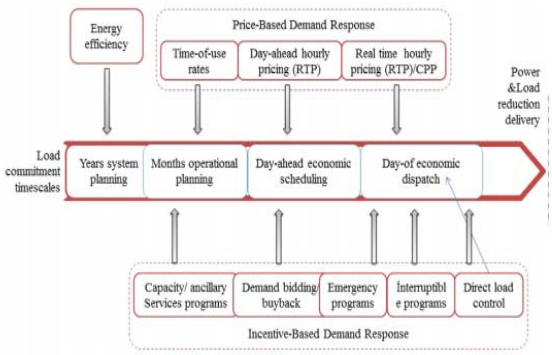


Fig. 2. Demand response in smart grid [12]

D. Employment of ''smart'' technologies as metering and distribution automation

Metering in smart grids supports two-way transmission between the meters and the utility. The meters guarantee further accurate bills and place customers in control of their energy usage. Smart meters include sensors, power outage notification, and power quality monitoring. Distribution automation is permanently linked with smart meters. Through advanced metering infrastructure (AMI), utilities can gather customer data more rapidly and offer a system-wide communications network to benefit service points and connect devices throughout the grid. The AMI and distribution automation unlocks the door for vast grid modernization yet transformer and feeder monitoring, outage management, integration of electric vehicles, and effectual fault isolation [13]. A unique technique to succeed distribution automation is within the employment of Substation Automation System (SAS) that outlines locally controlled actions to resolve congestion with minimal renewable energy resource limitations [14].

E. Integration of ''smart'' appliances and customer devices

Smart appliances and devices are parts of equipment that could connect with electrical grids, turn off throughout peak hours and change energy usage both intellectually and individually. In Britain's research, the demand response in households with a 20% perception of smart appliances can run up to 54% of the operating reserve necessities reliant on time of the day [15]. Smart appliances change household electricity demand. A wireless sensor domestic area network working on ZigBee protocol, utilized for transmitting messages between different units in a household energy management system established on appliances coordination, can offer improved resolutions for energy management problems [16].

F. Innovative electricity storage and peak-shaving technologies

Electricity storage and technologies that endeavor to adequately lower peaks is a crucial function of the smart grid. Energy storage is essential due to electricity generation from renewable energy oscillates. Storage devices keep the extra electricity when renewable energy generation is abundant thus the system is capable of utilizing this energy as demand enhances. Electric vehicles (EVs) can assist the electric grid as an individual energy resource. They can stay linked to the grid when they are parked, therefore distribute the energy from their batteries in a technology known as vehicle-to-grid (V2G) [17].

G. Timely information and control option

Timely information throughout the electrical grid is obtained when grids become smarter. Generation, transmission, distribution, and customers must make informed choices at the highest suitable time. Time synchronization and intellectual end-point devices permit the gathering of the necessary information for more rapid recognition of illegal customers, branch overload detection, and power-quality verification [18]. However, intellectual control in smart grids is essential for the ideal planning of

energy sources to increase power transport for transient stability and for real power control. It enhances power generation through perfectly tracing the load demand fluctuations [19] and can be worked for dependent error detection and reconfiguration [20]. Reddy et al. [21] underline the various methods for smart grid significant characteristics, Integration, Control, Communication and Metering (ICCM). Supplying precise timestamping of network events, of key importance in any time synchronization protocol, is a major disturb low-cost and low-complexity timestamping techniques that sustain full compatibility with already existing communication standards for wireless nodes used in smart grids suggests a solution.

H. Interoperability of appliances

Interoperability of appliances and equipment linked to the electric grid, involving the infrastructure helping the grid. Grid parts should be adept to run together to allow the reliable transportation of electricity from generation to expenditure. Interoperability is essential in designing and employing grid architecture. The complexity and the amount and scale of systems and devices engaged in smart grids make interoperability essential. The smart grid Interoperability Panel (SGIP), introduced by NIST is aimed to synchronize standards development for smart grids and verify that the parts can flawlessly work and communicate with each other; all integrated domains of the power system, consumers, markets, service providers, operations, bulk generation, transmission and distribution work together to create a modern, and efficient grid [22].

V. SMART GRID TECHNOLOGY AND APPLICATION

Smart grid conceptions involve an inclusive sort of technologies and applications. We illustrate a few below that are presently in practice with the limitation that are at this early step in the improvement of smart grids [23]:

A. Advanced metering infrastructure (AMI)

(AMI) is a concept for two-way meter/utility transmission. Two essential elements of AMI have been employed. First, automatic meter reading (AMR) systems deliver an early step toward reducing the costs of information gathering via the use of real-time metering information. Second, meter data management (MDM) delivers a single point of integration for the complete range of meter information. It permits leveraging that information to automate business processes in real-time and sharing of the information with key business and operational applications to enhance efficiency and support decision making across the enterprise.

B. Distribution management system (DMS)

(DMS) software arithmetically models the electric distribution network and expects the effect of outages, transmission, generation, voltage/frequency variation, and more. It services decreasing capital investment via demonstrating how to improve utilization existing assets, through enabling peak shaving via demand response (DR), and through improving network reliability.

C. Geographic information system (GIS)

(GIS) technology is exclusively created for the utility industry to model, design, and manage their significant infrastructure. Through integrating utility data and geographical maps, GIS delivers a graphical vision of the infrastructure that supports cost reduction throughout simplified planning.

VI. DATA SOURCES

There are different data categories according to the kind of extorted values: (i) Operational data that is the electrical data of the grid which represent real and reactive power currents, demand response capacity, voltage. (ii) Non-operational data is not associated to grid power but it indicates to master data, data on power quality and reliability. (iii) Meter usage data is an extra type of data related to power usage and demand values such as average, peak and time of the day. (iv) Event message data derives from smart grid device events like voltage loss/restoration and fault detection events. Lastly, (v) metadata, which is consumed to coordinate and interpret all the other types of data. All these data are gathered from numerous suppliers such as meters, sensors, devices and substations.

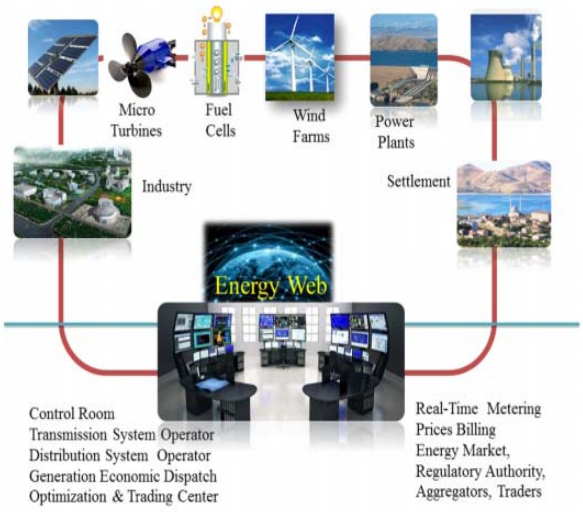


Fig. 3. Some of Smart Grid applications [24].

VII. DATA INTEGRATION

Latest data and communication technologies and advanced operation are depleted to enhance smart grid consistency, persistence, effectiveness and performance. The purpose of this is to have many technologies and methods to guarantee data integration:

A. *Service Oriented Architecture* (*SOA*)

All operation systems link a great number of software, each one having its specific and particular approach to offer services to consumers. Therefore, the question is how to handle and sustain all these systems. As a result, SOA builds software connected together using a single methodology that makes data integration simple and flexible [25].

B. *Enterprise Service Bus* (*ESB*)

Founded on a countless number of methods to manage communication between various types of systems such as GIS, ESB. ESB gives a lot of advantages to minimize cost and time in term of management, monitoring and divergence of integration [26]. In smart grids, ESB technologies are highly associated to SOA, making it more robust and flexible.

C. *Common Information Models* (*CIM*)

CIM is depleted for smart grid perseverance and for incorporated data construction and are crucial, particularly in the success or failure of data management. CIM denotes to UML models for the electric power manufacturing. It performs an obviously significant role in energy management systems in terms of data integration, time and cost. CIM help to switch data with technical grid structure. The CIM develop in power systems in order to assure the data interoperability, in the case of employing various applications.

D. *Messaging*

Stand for communication systems centered on exchanging messages. These messages involve data and other information from various applications managed by messaging server [26].

VIII. DATA STORAGE

Data storage in smart grids has a significant role, due to this it is based on gathering data from transmitted sources and delivering data to analytics tools in fast input/output operations per second (IOPS). Ultimately, there is a necessity for an advanced and accessible data storage procedure to encounter big data requirements.

A. *Distributed File System* (*DFS*)

A file system which permits several consumers on several machines to share documents and storage sources. It is built on client/server architecture as a storage device authorizing every consumer to receive a local manuscript of the stored data. There is a large amount of resolutions which use DFS such as: Google File System (GFS), Quant Cast File System (QCFS), Ceph (Linux-based distributed file system), Luster File System, Gluster File System.

B. *NoSQL databases*

Novel database method to permeate the limitations of traditional relational SQL databases in the situation of enormous data. These types of databases stage three constructions: key-value resolutions such as Dynamo and Voldemort, column-oriented resolutions such as Cassandra and HBase and document database resolutions such as MongoDB and CouchDB.

IX. DATA VISUILIZATION

Data visualization has an unlimited role because it adjusts the evaluation of smart grids. Visualization techniques are constructed on multivariate high dimensional visualization that yields the ability to deploy 2D and 3D visualization. However, smart grids encounter huge variables which complicate data presentation, such as 3D Power-map.

X. DATA TRANSMISSION

Data transmission in big data affects all the previous stages. Accordingly, it must sustain high bandwidth capacity and speed, data security and privacy. Data transmission in smart grids is constructed on communication technologies as described in "communication systems," beginning with access network technologies involving PLC (Programmable Logic Controller), ZigBee and WIFI, followed by area network technologies using M2M, cellular networks and Ethernet; afterward core network technologies with internet protocol address (IP) and Multiprotocol Label Switching (MPLS). Lastly, backbone network technologies, that depend on fiber technologies, microwave link, IP-based Wavelength, Division Multiplexing (WDM) network and other optical technologies.

XI. BIG DATA TOOLS FOR CUSTOMER DATA ANALYTICS

Customers' data is in the order of terabytes and in a diversity of structures which entails high velocity, scalability and fault tolerance in data processing, storage and visualization. Big data employment could be performed using many tools. Figure 4 shows many big data technologies which could be deployed to manage smart grid data. The diversity of consumer data sources (smart meters, devices, historical data) entails the use of integration tools to make data uniform. Messaging tools are the most proficient for raw data integration and therefore could be used for consumer data integration. Big data analytics can be done by deploying many processing modes:

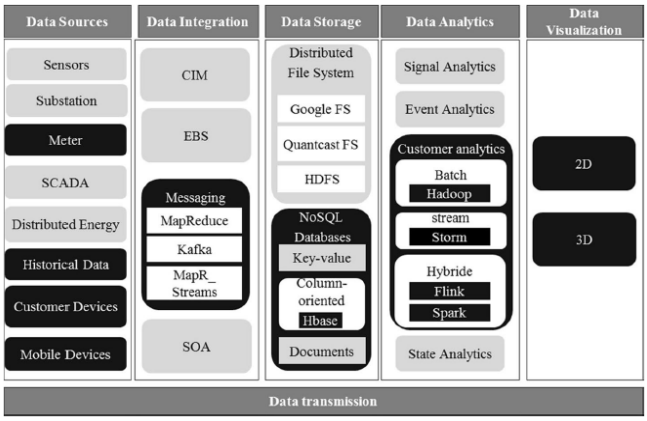


Fig. 4. Big data tools depending on the targeted customer data analytics

A. Batch processing tools

Big data analytics propose a large number of approaches to process data beginning from batch processing. Hadoop [27] is a proper choice for batch analytics for smart grids. While smart grid systems are delivered geographically, delivered file systems are extremely valuable for it. Hadoop has Hbase as a database system, Hadoop Distributed File System (HDFS) as a storage system, and MapReduce as a processing engine. Hadoop cannot control advanced Information Technology (IT) systems in data velocity, scalability and machine learning algorithms.

B. Real time processing tools

Real time processing is rapid in terms of performance, more so than batch processing due to its ability of controlling data with high velocity obligations using stream processing or complex event processing systems. Real time processing could be employed using some resolutions such as S4, Spelunk, and Storm. Storm [28] is an open source that offers fault-tolerance, suggesting a great number of prospects as a real time processing system, incorporating message handling reliability, parallel computations and simple programming models. Storm can be deployed with Kafka for data integration and Hbase for data storage.

C. Hybrid processing tools

Hybrid processing is able to manage both batch and real time processing. Spark [29] is a framework used for batch processing, but it has also real time processing resolution with Spark streaming. Spark manages large-scale data processing incorporating useful tools such as Spark SQL, Spark Streaming, machine learning libraries and GraphX. All that yield Spark meet big data obligations in smart grid. Spark streaming deploys a real time complex event processing engine to manage velocity problems. Whendeploying Spark, data storage could be completed using HDFS or even Hbase. Apache Flink [30] is a different framework thatcan process data in both batch and stream modes. Flink is centered on huge APIs like transformation functions which make it accessible, effortless to employ, error tolerant and rapid in execution. Flink is useful in machine learning due to its ability to adopt its particular machine learning library called FlinkML.

CONCLUSIONS

Smart grids are now an utmost inclusive technology, which have been developing quickly through recent years due to their advantages. Smart grid systems gather enormous amount of datasets which in turn develop more smart capabilities for the grid. However, this creates challenges for utilities to handle the nature, the delivery and the real-time restraints of the gathered data. In this paper we have staged an overview of the prospects, concepts and challenges of data management in smart grids and briefly described the big data technologies and procedures which could be deployed to manage smart grid obligations involving processing, storage and visualization. We have also explained the phases, tools and technical obligations for deploying and implementing big data technologies for smart grids in order to have effectual and accessible data management. The resulting evolvement of smart grids relies on the availability of sensitive and reliable measurement data for monitoring and control of these grids. To create a system in the smart grid system, special components must be molded to observe the voltage, the frequency, the harmonics, the current limits stated in energy and the power cuts made in the form of monitoring. The exact future of smart grids could be challenging, however current inventions present a dynamic merging of sectors, mechanics and communities.

REFERENCES

[1] Marco Liserre, Thilo Sauter, and John Y. Hung, “Future Energy Systems”, IEEE Industrial Electronics Magazine, March 2010.
[2] Cunjiang Y, Huaxun Z, Lei Z. Architecture design for smart grid. Energy Procedia 2012;17:1524–8. http://dx.doi.org/10.1016/j.egypro.2012.02.276
[3] Güngör VC, Sahin D, Kocak T, Ergüt S, Buccella C, Cecati C, Hancke GP. Smart grid technologies: communication technologies and standards. IEEE Trans Ind Inform 2011;7(4):529–39.
[4] Congress of the United States of America. Energy independence and security act of 2007 [Public Law 110–140]; 2007.
[5] Doguc Ozge, Emmanuel Ramirez-Marquez Jose. An automated method for estimating reliability of grid systems using Bayesian networks. Reliab Eng Syst Saf 2012;104:96–105.
[6] Colmenar-santos A, Pérez M-ángel, Borge-diez D, Pérez-molina C. Reliability and management of isolated smart-grid with dual mode in remote places : application in the scope of great energetic needs. Electr Power Energy Syst 2015;73:805–18. http://dx.doi.org/10.1016/j.ijepes.2015.06.007.
[7] Suleiman H, Alqassem I, Diabat A, Arnautovic E, Svetinovic D. Integrated smart grid systems security threat model. Inf Syst 2015;53:147–60. http://dx. doi.org/10.1016/j.is.2014.12.002.
[8] Electric Power Research Institute. Distributed energy resources 2015. ⟨http:// www.epri.com/Our-Work/Pages/Distributed-Electricity-Resources.aspx⟩ [accessed 25.9.17].
[9] Howlader HOR, Matayoshi H, Senjyu T. Distributed generation incorporated with the thermal generation for optimum operation of a smart

grid considering forecast error. Energy Convers Manag 2015;96:303–14. http://dx. doi.org/10.1016/j.enconman.2015.02.087.
[10] Federal Energy Regulatory Commission. Reports on demand response & advanced metering n.d. ⟨https://www.ferc.gov/industries/electric/indus-act/ demand-response/dem-res-adv-metering.asp⟩ [accessed 25.10.17].
[11] U.S. Department of Energy. Demand response n.d. ⟨http://energy.gov/oe/ technology-development/smart-grid/demand-response⟩ [accessed 25.10.17].
[12] Wissner M. The Smart Grid - a saucerful of secrets? Appl Energy 2011;88 (7):2509–18.
[13] Hart DG. How advanced metering can contribute to distribution automation. IEEE Smart Grid. ⟨http://smartgrid.ieee.org/august-2012/644-how-advancedmetering-can-contribute-to-distribution-automation⟩; 2012 [accessed 26.07.17].
[14] Berizzi A, Bovo C, Ilea V, Merlo M, Miotti A, Zanellini F. Decentralized congestion mitigation in HV distribution grids with large penetration of renewable generation. Int J Electr Power Energy Syst 2015;71:51–9. http: //dx.doi.org/10.1016/j.ijepes.2015.02.023.
[15] Nistor S, Wu J, Sooriyabandara M, Ekanayake J. Capability of smart appliances to provide reserve services. Appl Energy 2015;138:590–7. http://dx.doi.org/ 10.1016/j.apenergy.2014.09.011.
[16] Mahmood A, Khan I, Razzaq S, Najam Z, Khan NA, Rehman MA, et al. Home appliances coordination scheme for energy management ((HACS4EM)HACS4EM using wireless sensor networks in smart grids. Procedia Comput Sci 2014;32:469–76. http://dx.doi.org/10.1016/j.procs.2014.05.449.
[17] University of Delaware. Car prototype generates electricity, and cash – sciencedaily. Sciencedaily n.d. ⟨http://www.sciencedaily.com/releases/2007/12/ 071203133532.htm⟩ [accessed 29.11.17].
[18] Stastny L, Franek L, Bradac Z. Time synchronized low-voltage measurements for smart grids. Procedia Eng 2015;100:1389–95. http://dx.doi.org/10.1016/j. proeng.2015.01.508.
[19] Yilmaz M, Dhansri NR. A smart grid intelligent control framework. IEEE Green Technol. Conf 2012;3(19–20):1. http://dx.doi.org/10.1109/GREEN.2012.6200983.
[20] Tarhuni NG, Elkalashy NI, Kawady TA, Lehtonen M. Autonomous control strategy for fault management in distribution networks. Electr Power Syst Res 2015;121:252–9. http://dx.doi.org/10.1016/j.epsr.2014.11.011.
[21] Reddy KS, Kumar M, Mallick TK, Sharon H, Lokeswaran S. A review of integration, control, communication and metering (ICCM) of renewable energy based smart grid. Renew Sustain Energy Rev 2014;38:180–92.
[22] Smart Grid Interoperability Panel. About SGIP 2015. ⟨http://sgip.org/AboutSGIP⟩ [accessed 29.11.17].
[23] T.Samad and A.M. Annaswamy, "The Impact of control technology-Control for renewable energy and Smart Grid" www.ieeecss.org. (eds), 2011.
[24] Yan Y, Qian Y, Sharif H, Tipper D. A survey on smart grid communication infrastructures: Motivations requirements and challenges. IEEE Commun Surv Tutor 2013;15(1):5–20.
[25] Minguez J, Jakob M, Heinkel U. A soa-based approach for the integration of a data propagation system. In: Proceedings of the 9th international conference on information reuse and integration. 10–12 August 2009. Melbourne; 2009. p 47–52
[26] Vera-Baquero A, Colomo-Palacios R, Molloy O. Business process analytics using a big data approach. IT Prof. 2013;15:29–35.
[27] Messaging Integration. http://www.informationbuilders.com/messaging
[28] Apache Hadoop. http://hadoop.apache.org. Accessed 29 Jan 2017
[29] Apache Storm. http://storm.apache.org. Accessed 5 Feb 2017
[30] Apache Spark. http://spark.apache.org. Accessed 20 Feb 2017
[31] Apache Flink. https://flink.apache.org. Accessed 15 Jan 2017

On-Body Humidity Sensing Antenna with Polyimide for BAN Applications over 5G Networks

Md. Majedur Rahman
Department of EEE
American International University-Bangladesh (AIUB)
Dhaka, Bangladesh
majedur-rahman@outlook.com

Md. Abu Nayeem
Department of EEE
American International University-Bangladesh (AIUB)
Dhaka, Bangladesh
abu.nayeem@outlook.com

Saqlain Nahid
Department of EEE
American International University-Bangladesh (AIUB)
Dhaka, Bangladesh
nahid.saqlain.bd@gmail.com

Syed Raqueed Bin Alvee
Department of EEE
American International University-Bangladesh (AIUB)
Dhaka, Bangladesh
alveeraqueed@gmail.com

Raja Rashidul Hasan
Department of EEE
American International University-Bangladesh (AIUB)
Dhaka, Bangladesh
hemal@aiub.edu

Md. Abdur Rahman
Department of EEE
American International University-Bangladesh (AIUB)
Dhaka, Bangladesh
arahman@aiub.edu

***Abstract*—This paper proposes an on-body humidity sensing antenna with polyimide working in the 5G network. The chosen operating frequency is 38 GHz which also lies in the mm-wave band. This paper discusses two antennas. The first antenna is designed using polyimide film and the other using polyimide film as a superstate with Rogers RT 5880 as the main substrate. The first antenna exhibits an intensive radiation absorption of 38.7 W/kg for every 10g of tissue, which is mitigated by the design of the second antenna. Therefore, the second antenna is analyzed for on-body humidity sensing. Due to polyimide's high sensitivity towards humidity, any change in humidity is detectable through the changes in the dielectric constant of polyimide and changes in the resonant frequency.**

***Index Terms*—on body, microstrip antenna, humidity sensing antenna, millimeter wave, 5G, wearable antenna, BAN**

I. Introduction

Wireless communication has been around only for a few decades, yet it has escalated in its field in remarkable ways. Reuse of frequencies in an effective way, the introduction of digital modulation, packet-based internet penetration, and enormous enhancement in physical layer technologies as in WCDMA, OFDMA, MIMO, HARQ, etc. comprehensively assisted toward this escalation [1]. 5G technology is the latest technology and the culmination of all the precedent wireless communications technologies in terms of speed, latency, bandwidth, and energy consumption [2]. In microwave frequency, the spectrum has become more elusive, unlike the millimeter-wave frequency [12]. In this paper, the operating frequency is selected from the millimeter-wave band which has a whole range of the unexploited spectrum (3 -300 GHz) [3]. But for the 5G applications, only the spectrum of 20 – 90 GHz is feasible [4]. From the available spectrum, 38 GHz was chosen as the center frequency which lies in the Ka-band (27-40 GHz) as well [5]. Frequencies in this band promise to have a high data rate and low latency system [6].

Body Area Network (BAN) is beginning to become a lucrative field of interest for researchers and the quantity of researches is being carried out in this field is comprehensive. They are beginning to become crucial in a broad scale of applications from healthcare to entertainment [7]. BAN is defined as the prominent wireless sensor network (WSN) which is planted on the human body, endorsing the physiological sensors possessing intelligent processing and wireless communication functionality having one motive of observing parameters such as body temperatures, skin surface humidity, respiratory rate, and blood pressure [8].

5G possesses no scarcity and provides true wireless communication to society. Apart from that, this technology will come at much low cost and will ensure high peak expectations comparing to its precedent technologies. 5G is alleged to bring the perfect wireless world or called "WWWW: World Wide Wireless Web" [11]. The combination of 5G antenna working in the Ka-band with BAN will bring enrichment to the motive of this paper.

In this paper the primary dielectric material is polyimide. Polyimide is a kind of heteroaromatic polymer [9]. For its wide range of applications, it provides the scope of exploitation to the researchers. It has high-temperature tolerance and a broad variety of applications as mechanical parts, electronic parts, electrical insulation, pressure-sensitive tape, fiber optics cable, insulation blankets, insulation tubing, automotive diaphragms sensors, and manifolds [10]. Along

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

with all these applications it is also used in humidity sensing, some of the works are discussed in the next section.

The remaining of the paper is arranged as follows. Section II is discussed with some related work in humidity sensing with and without polyimide. In section III the design model of the two antennas is briefed, one with polyimide as the main substrate and the other with polyimide as a superstate. In section IV the simulations and results are discussed for both of the antennas and then compared with each other. Finally in section V brings the discussions about some of the flaws of the antennas and the intended future work for this research.

II. PREVIOUS WORKS

Polyimide has been around for a while and the research interest on polyimide as a humidity sensor also primitive. In [13], a humidity sensor's design is studied which is based on the change of relative permittivity of polyimide. They used the thin film of polyimide which was claimed to be the key component of the sensor. The work mentioned the construction of a capacitive structure in which the average dielectric constant changed through the diffusion of water through polyimide.

Polyimide is also found as a layer on top of the patch of the antenna for a different purpose and humidity sensing is also not an exception. For example, in [15] polyimide is used as a superstate for humidity sensing. They constructed a capacitive sensor with a split ring resonator (SRR). Later the resonator is coated on the top with polyimide film for humidity sensing. The use of polyimide as a superstate can bring escalating enhancement to an antenna, like better performance, realized gain, directivity, return loss, even it can be used as a frequency changer that can fill the motive of a definite purpose [14].

There are also works for humidity sensing based on RFID tags. In [16], they presented a prototype of an RFID humidity sensor based on a passive RFID tag. The sensing is done through the far-field backscattering coupling which uses electromagnetic backscattering for communication. They painted the surface of the RFID tag with polyimide which is a humidity sensing material itself. Besides in [17], there is another passive RFID system that is inkjet printed, designed to sense humidity. It is a passive UHF RFID system. This paper also exploits the change in frequency concerning the dielectric constant for better observations.

Humidity sensing for on body, wearable and, BAN devices is very elusive. But if the fabric antenna for humidity sensing is considered, there are some works. For example, in [18] they used inkjet technology for printing humidity sensors on textiles. The main application they were looking for was in smart wearable electronics, as in smart shirts or socks. In [19] they developed a wearable antenna for BAN applications which will sense humidity. They also worked in 5G, millimeter wave spectrum and, used cotton fabric as the main substrate of their antenna. They made it conspicuous that, humidity has a comprehensive impact on frequency and frequency could be an ideal parameter for the observation of the change in humidity.

III. ANTENNA DESIGN

The key motive of our paper is to design an antenna for on-body devices, hence it will be wearable and will sense humidity. It is quite vivid that, polyimide is a very promising candidate. In this paper, we designed two microstrip patch antenna, one using polyimide film as the main substrate and the other where polyimide film is in superstate as in [14], [15] with Rogers RT 5880 as the main substrate. For both of the antenna, the chosen center frequency is 38 GHz. Apart from the mm-wave band this frequency also lies to Ka-band [5]. Besides this specific center frequency gives more throughput performances and higher bandwidth compared to its other frequencies in its band [2].

For simulation and designing, Computer Simulation Studio (CST) was used. For the first antenna as shown in Figure 1, polyimide film, Kapton was used ($\epsilon_r = 3.5$, tanδ = 0.0027, h = 0.25 mm) as the substrate [10]. For the second antenna as shown in Figure 2, Rogers RT 5880 is used ($\epsilon_r = 2.2$, tanδ = 0.0009, h = 0.254 mm) as the substrate. As a superstate, 100 μm of polyimide film was used on top of the patch of the antenna. For conductive parts (antenna patch, ground plane and transmission line), copper ($\sigma = 5.8E + 007S/m$) was used, which is about 35 μm thick.

The width (W) and length (L) of the patch of an antenna plays a crucial part in the design. They are determined as

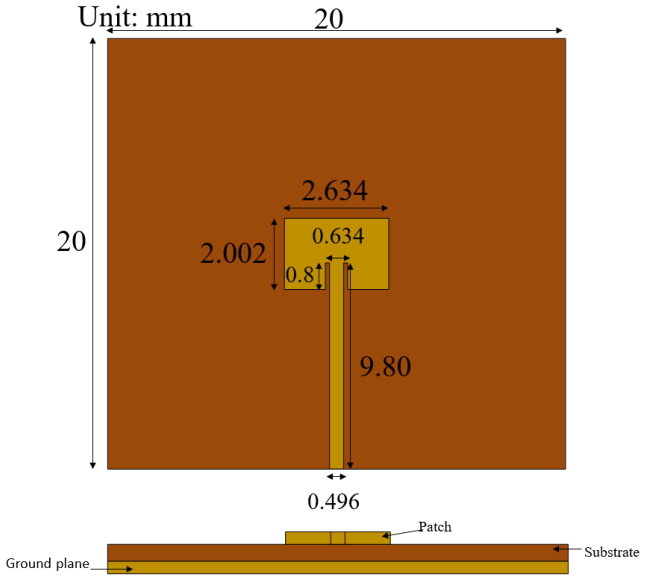


Fig. 1: The geometry of the proposed antenna with polyimide as a substrate

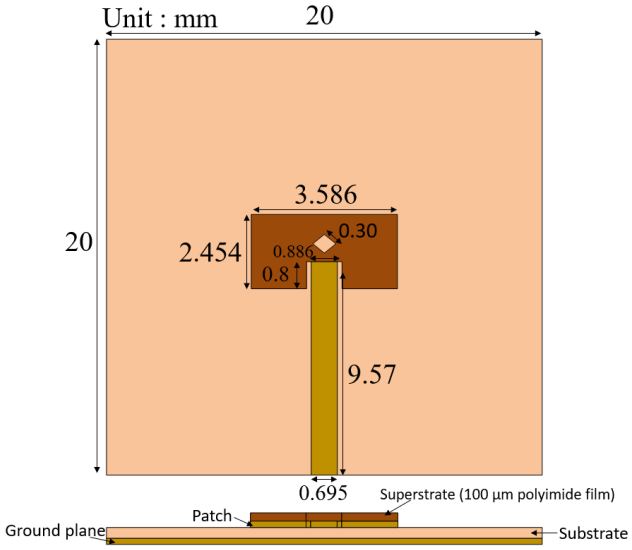


Fig. 2: The geometry of the proposed antenna with polyimide as a superstate

following equations 1 to 4 [20].

$$W = \frac{c}{2f_r}\sqrt{\frac{2}{\epsilon_r + 1}} \tag{1}$$

where c is the speed of light in free space and f_r is desired resonant frequency at which the antenna operates.

$$L = \frac{c}{2f_r\sqrt{\epsilon_{reff}}} - 2\Delta L \tag{2}$$

where ϵ_{reff} is the effective dielectric constant of the substrate and ΔL is extended length due to fringing effect.

$$\epsilon_{reff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2}[1 + \frac{12h}{W}]^{-\frac{1}{2}} \tag{3}$$

$$\Delta L = 0.412h\frac{(\epsilon_{reff} + 0.3)(\frac{w}{h} + 0.264)}{(\epsilon_{reff} - 0.258)(\frac{w}{h} + 0.8)} \tag{4}$$

where h is substrate thickness.

In order to match the 50-Ω characteristic impedance of coaxial cable [21], with the trace width of the antenna (W_f) following equation is followed.

$$Z_o = \frac{120\pi}{\sqrt{\epsilon_{reff}}}\frac{1}{[\frac{W_f}{h} + 1.393 + 0.677 \times \ln(\frac{W_f}{h} + 1.444)]} \tag{5}$$

where Z_o = 50-Ω characterized feed line impedance.

IV. Simulation and Result Analysis

A. Polyimide as Substrate

The first antenna as shown in Figure 1, was designed only with a polyimide film. Since polyimide is very sensitive to humidity, small changes in humidity could be observed through the small changes of the dielectric constant of the polyimide [13], [15].

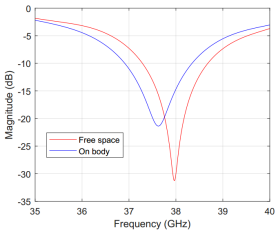


Fig. 3: Reflection coefficient (S11) of the antenna with polyimide as a substrate

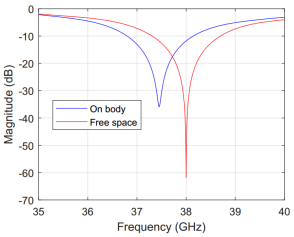


Fig. 4: Reflection coefficient (S11) of the antenna with polyimide as a superstate

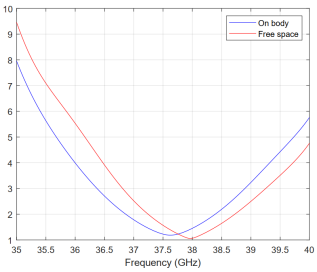


Fig. 5: VSWR of the antenna with polyimide as a substrate

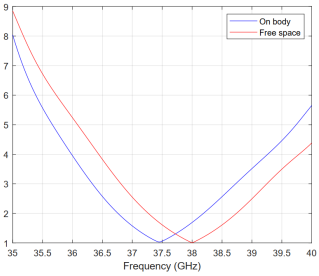


Fig. 6: VSWR of the antenna with polyimide as a superstate

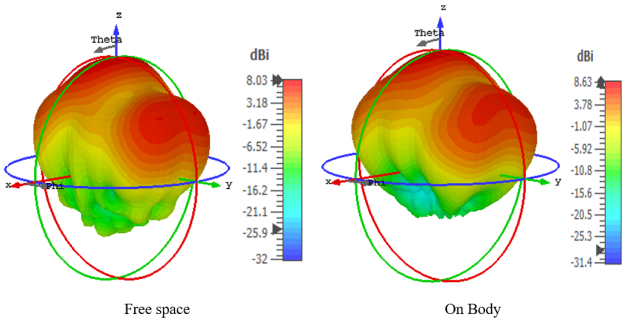


Fig. 7: Far-field directivity of the antenna with polyimide as a substrate

The results of the antenna as shown in Figure 3, exhibits that, in free space, it has a reflection coefficient of -31.25 dB, which is way below the rule of thumb of -10 dB with the resonant frequency of 37.965 GHz. The bandwidth of the antenna is 1.394 GHz with a lower and higher frequency of 37.286 GHz and 38.68 GHz respectively. The fractional bandwidth is 3.67%, therefore a deviation from center frequency within this range will be functional for the antenna. On phantom, the resonant frequency reduced to 37.625 GHz with a reflection coefficient of -21.37 dB, which still lies within the bandwidth of the antenna. Now from Figure 5, it can be seen that the voltage standing wave ratio (VSWR) of the antenna is 1.056 in free space, while on the body the VSWR raised to 1.187. The values of VSWR, both fall in the range of 1-2, implying proper impedance matching. Observation from Figure 7 shows

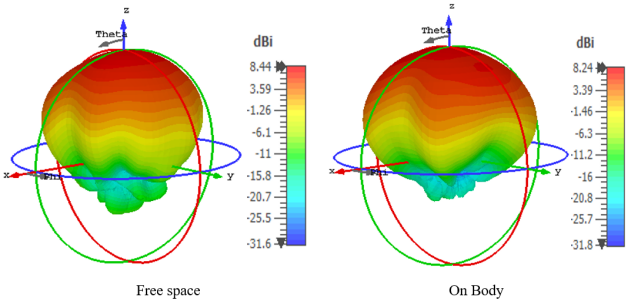


Fig. 8: Far-field directivity of the antenna with polyimide as a superstate

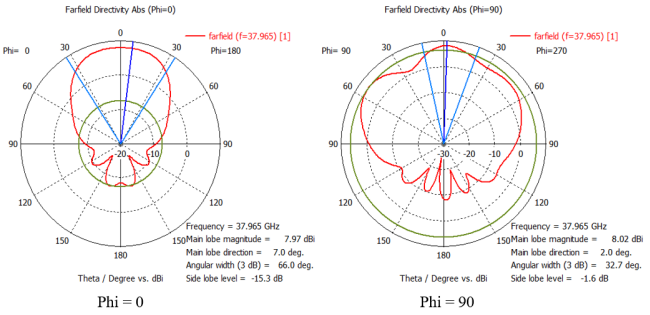


Fig. 9: Radiation pattern of the antenna with polyimide as a substrate

that the antenna results in the directivity of 8.02 dBi with a radiation efficiency of -1.293 dB in free space which is approximately 74.25%. On-body, the directivity of the antenna is 8.63 dBi. Figure 9 shows the radiation pattern of the antenna at phi = 90 and phi= 0.

B. Polyimide as Superstate

The second antenna as shown in Figure 2, was designed with polyimide as a superstate. In this design, the polyimide film was intended to use as a top layer on the patch of the antenna. As the main substrate, Rogers RT 5880 was used, as it has a good plausibility of having better efficiency, reflection coefficient, and a better Specific Absorption Rate [22].

The results as shown in Figure 4 exhibit that, in free space, the antenna has a very fitting reflection coefficient of -61.82 dB with a resonant frequency of 38 GHz. The bandwidth of the antenna is 1.389 GHz with a lower and higher frequency of 37.309 GHz and 38.698 GHz respectively. The fractional bandwidth is 3.65 %, hence a deviation within this range is feasible for antenna functionality. On-body phantom, the resonant frequency reduced to 37.45 GHz with a reflection coefficient of -35.88 dB, which still lies within the bandwidth of the antenna. Now from Figure 6, it can be seen that the VSWR of the antenna is 1.001 in free space, while on-body the VSWR raised to 1.033. From Figure 8, the antenna results

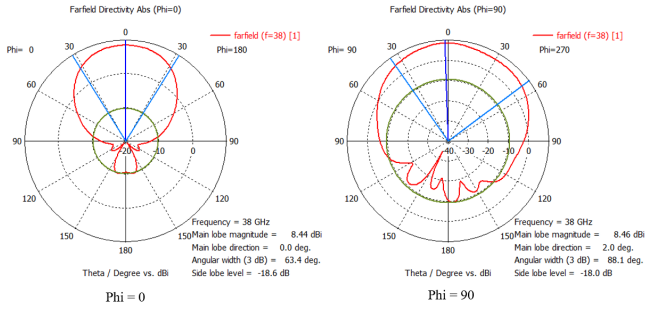


Fig. 10: Radiation pattern of the antenna with polyimide as a superstate

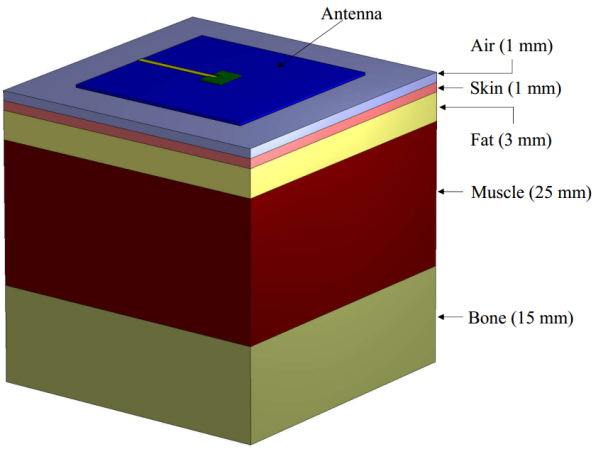


Fig. 11: Model of the body phantom in CST

in the directivity of 8.46 dBi with a radiation efficiency of -0.9634 dB in free space which is approximately 80.11%. On-body the directivity of the antenna is 8.24 dBi. Figure 10 shows the radiation pattern of the antenna at phi = 90 and phi= 0.

C. Specific Absorption Rate (SAR) Analysis

Specific Absorption Rate is a measure of the rate at which an object especially human tissue absorbs energy when exposed to the electromagnetic field [26].To protect the human body from the harmful radiation, the International Commission on Non-Ionizing Radiation Protection (ICNIRP) demarcated the safety limit for SAR. According to ICNIRP for every 10g of human tissue, the SAR shouldn't surpass 2.0 W/Kg [24]. Besides the Federal Communications Commission (FCC) made the safety protocol more subtle by appointing the limit for SAR 1.6 W/Kg for every 1g of human tissue [25]. The antenna which works for BAN applications or very close to the human body, they need to satisfy the SAR analysis. The proposed antennas in this paper are intended to work from on top of the human body. Hence it is also important for them to perform well while they are in close

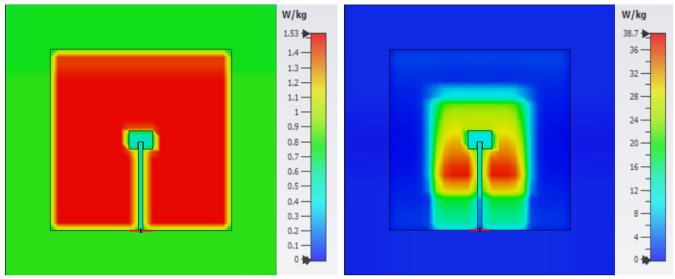
contact with the human body.

The human body is lossy and hence always the results on the human body deviate from the results taken in free space. We have already seen such deviations for both of the antennas above. The dielectric properties as shown in Table I, were followed for creating the body phantom in CST (see Figure 11) and to execute the simulations especially to test the SAR.

The first antenna where polyimide film was used as the main substrate (see Figure 1), showing an elevated value of SAR. As shown in Figure 12, for every 10g of tissue it absorbed a SAR of 1.53 W/Kg and 38.7 W/Kg for every 1g of tissue. The second antenna where the polyimide film was used as a superstate (see Figure 2), showed a rather less absorption of SAR. As in Figure 13, for every 10g of tissue, it absorbed a SAR of 0.41 W/Kg and 1.59 W/Kg for every 1g of tissue.

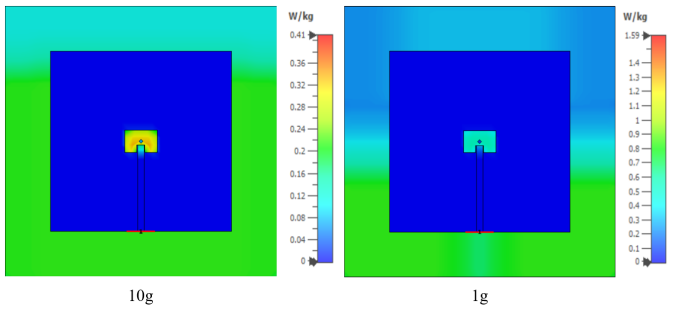
TABLE I: Dielectric Properties of Human Tissues at 38 GHz [23]

	Dielectric constant ϵ_r	***Loss tangent*** $tan\delta$	***Conductivity*** σ***(S/m)***
Dry skin	12.297	1.1941	31.043
Fat	3.4444	0.29331	2.1358
Muscle	19.056	1.0382	41.823
Bone cortical	4.5241	0.61275	5.8603



10g 1g

Fig. 12: SAR values for the proposed antenna with polyimide as a substrate



10g 1g

Fig. 13: SAR values for the proposed antenna with polyimide as a superstate

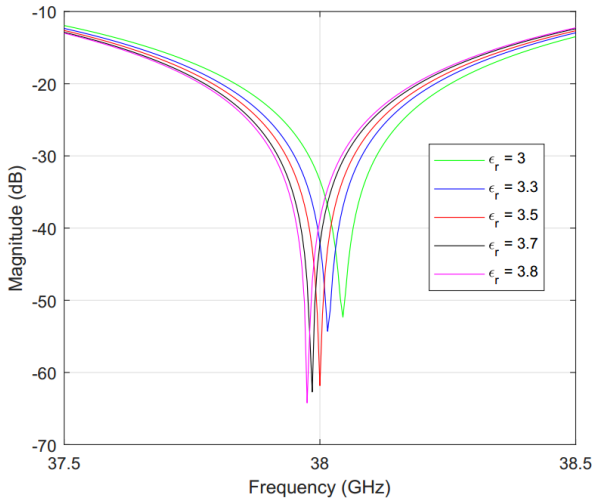


Fig. 14: Change of operating frequency concerning the relative permittivity of polyimide film for humidity sensing

TABLE II: Summary for the Change in Resonant Frequency for Humidity Sensing

Dielectric constant ϵ_r	***Operating frequency (GHz)***	***Reflection coefficient (dB)***	***Relative humidity (%)***
3.0	38.045	-52.34	0
3.3	38.015	-54.29	30
3.5	38	-61.82	50
3.7	37.985	-62.69	80
3.8	37.975	-64.22	100

D. Humidity Sensing

The SAR analysis of the first antenna showed a high elevation on radiation absorption which is not acceptable for antenna working nearby of the human body. Although it had been resuscitated by eliminating polyimide film as substrate and replacing it with Rogers RT 5880 which exhibits better SAR [22]. Hence the design of the second antenna tested further for the observation of humidity sensing. Now for humidity sensing based on the change in the demeanor of the dielectric material, the relative permittivity of the polyimide film had been observed based on the datasheet of polyimide film Kapton [27]. From Figure 14, the overall effect of humidity on the operating frequency of the antenna can be seen in a brief range. For a subtle observation the data of Figure 14 was extracted in Table II.

Now if the values of Table II are considered, it is quite vivid that the relation of the humidity and frequency is not linear as operating frequency decreased with the rise in relative humidity. As the dielectric constant of polyimide increases with relative humidity [27]. For each in-between steps, the frequency shift was around 10-30 MHz. The relative permittivity of polyimide film, ϵ_r = 3.5 is the base value for this design as the antenna was designed at that value. In the beginning, the

operating frequency of the antenna followed a pattern. For the first upward and downward step when ϵ_r = 3.7 and ϵ_r = 3.3 respectively, the resonant frequency had a consistent change of 15 MHz. But when relative humidity moves to 0% and 100%, the change in the resonant frequency become inconsistent, with an increment of 45 MHz and decrement of 25 MHz in resonant frequency respectively. The values of the reflection coefficient justify that, the antenna will work in all of the mentioned values of relative humidity as shown in Table II.

V. Conclusion

The main goal of this paper is to design an antenna that will sense humidity over the 5G network while remaining close to the human body. We presented two antennas both designed in CST working at 38 GHz which lies in the mm-wave band. The first antenna is designed using only polyimide film as a substrate. It exhibited quite good results with good bandwidth, gain, and return loss. But it had a high radiation absorption rate for both 10g and 1g of tissue. Which later recuperated by replacing the main substrate with Rogers RT 5880 and using polyimide film only in a superstate. The second antenna brought quite promising results in terms of bandwidth, gain, return loss, and SAR. Hence this antenna later analyzed for humidity sensing where it also brings comprehensive results. The change in resonant frequency concerning the change of the relative permittivity of polyimide film was also quite conspicuous. This antenna will be able to sense humidity as any change in humidity will be easily observed through the change of the relative permittivity of polyimide film and change of resonant frequency which will be in the megahertz range.

In the future, we would like to suggest more antenna work based entirely on polyimide. Polyimide is a very sensitive material. There could be a lot of plausibility for polyimide if the radiation problem is eliminated. Moreover, there could be work on our current antenna in a different band of mm-wave like 73 GHz. Besides we would like to suggest more improvement on our current polyimide superstate antenna by reducing the gap between on body and free space operating frequency difference, which will make the antenna more accurate on its performances. Besides, we also would like to suggest the design of novel microstrip antennas for multiple sensing.

References

[1] M. Agiwal, A. Roy and N. Saxena, "Next Generation 5G Wireless Networks: A Comprehensive Survey," in IEEE Communications Surveys & Tutorials, vol. 18, no. 3, pp. 1617-1655, thirdquarter 2016.

[2] F. Qamar, M. H. S. Siddiqui, K. Dimyati, K. A. B. Noordin and M. B. Majed, "Channel characterization of 28 and 38 GHz MM-wave frequency band spectrum for the future 5G network," 2017 IEEE 15th Student Conference on Research and Development (SCOReD), Putrajaya, 2017, pp. 291-296.

[3] D. Imran et al., "Millimeter wave microstrip patch antenna for 5G mobile communication," 2018 International Conference on Engineering and Emerging Technologies (ICEET), Lahore, 2018, pp. 1-6.

[4] N. Al-Falahy and O. Y. K. Alani, "Design considerations of ultra dense 5G network in millimetre wave band," 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN), Milan, 2017, pp. 141-146.

[5] R. K. Gupta, T. Shanmuganantham and R. Kiruthika, "A staircase hexagonal shaped microstrip patch antenna for multiband applications," 2016 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), Kumaracoil, 2016, pp. 298-303.

[6] W. Ahmad and W. T. Khan, "Small form factor dual band (28/38 GHz) PIFA antenna for 5G applications," 2017 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM), Nagoya, 2017, pp. 21-24.

[7] M. Seiffert, F. Holstein, R. Schlosser and J. Schiller, "Next Generation Cooperative Wearables: Generalized Activity Assessment Computed Fully Distributed Within a Wireless Body Area Network," in IEEE Access, vol. 5, pp. 16793-16807, 2017.

[8] M. U. Talha and J. Ahmad, "Body area networks (BANs)–An overview with smart sensors based telemedical monitoring system". International Journal of Computer Applications, vol. 84, no. 8, 2013, pp. 20–23, 2013.

[9] W. W. Wright and M. Hallden-Abberton, "Polyimides," Ullmanns Encyclopedia of Industrial Chemistry, 2000.

[10] "Kapton® HN General-Purpose Polyimide Film," Dupont Electronic Solutions. [Online]. Available: https://bit.ly/3iDbGp4. [Accessed: 31-Mar-2020].

[11] A. Gohil, H. Modi, and S. K. Patel, "5G technology of mobile communication: A survey," 2013 International Conference on Intelligent Systems and Signal Processing (ISSP), 2013.

[12] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," IEEE Communications Magazine, vol. 52, no. 2, pp. 74–80, 2014.

[13] P. Schubert and J. Nevin, "A polyimide-based capacitive humidity sensor," IEEE Transactions on Electron Devices, vol. 32, no. 7, pp. 1220–1223, 1985.

[14] A. Dahalan, M. Aziz, M. Othman, H. Nornikman, and S. Ibrahim, "Kapton Based Antenna Simulation Analysis for Smart System," International Journal of Integrated Engineering, vol. 11, no. 4, May 2019.

[15] S. A. M. Akhir, S. Z. Ibrahim, N. Rosli, A. S. M. Zain, and N. Khalid, "Antenna for humidity sensor using split ring resonator," Indonesian Journal of Electrical Engineering and Computer Science, vol. 13, no. 2, p. 584, Jan. 2019.

[16] Y. Jia, M. Heiß, Q. Fu, and N. A. Gay, "A Prototype RFID Humidity Sensor for Built Environment Monitoring," 2008 International Workshop on Education Technology and Training & 2008 International Workshop on Geoscience and Remote Sensing, 2008.

[17] J. Virtanen, L. Ukkonen, T. Bjorninen, A. Z. Elsherbeni, and L. Sydänheimo, "Inkjet-Printed Humidity Sensor for Passive UHF RFID Systems," IEEE Transactions on Instrumentation and Measurement, vol. 60, no. 8, pp. 2768–2777, 2011.

[18] J. Weremczuk, G. Tarapata, and R. Jachowicz, "Humidity Sensor Printed on Textile with Use of Ink-Jet Technology," Procedia Engineering, vol. 47, pp. 1366–1369, 2012.

[19] X. Lin, B.-C. Seet, and F. Joseph, "Wearable humidity sensing antenna for BAN applications over 5G networks," 2018 IEEE 19th Wireless and Microwave Technology Conference (WAMICON), 2018.

[20] C. A. Balanis. Antenna theory: analysis and design. Hoboken, NJ: John Wiley & Sons, 2012.

[21] P. Emeraldi and U. Khayam, "Design of matching impedance for ultra wideband partial discharge detection," 2013 International Conference on Information Technology and Electrical Engineering (ICITEE), 2013.

[22] N. O. Parchin, M. Shen, and G. F. Pedersen, "Small-Size Tapered Slot Antenna (TSA) Design for use in 5G Phased Array Applications," Applied Computational Electromagnetics Society Journal, vol. 32, no. 3, pp. 193–202, Mar. 2017.

[23] "Dielectric Properties of Body Tissues in the frequency range 10 Hz - 100 GHz," ITALIAN NATIONAL RESEARCH COUNCIL Institute for Applied Physics. [Online]. Available: http://niremf.ifac.cnr.it/tissprop/htmlclie/htmlclie.php. [Accessed: 04-Apr-2020].

[24] "ICNIRP GUIDELINES FOR LIMITING EXPOSURE TO ELECTROMAGNETIC FIELDS (100 KHZ TO 300 GHZ)," ICNIRP, 2020. [Online]. Available: https://www.icnirp.org/cms/upload/publications/ICNIRPrfgdl2020.pdf. [Accessed: 22-Jun-2020].

[25] "Specific Absorption Rate (SAR) for Cellular Telephones," Federal Communications Commission, 22-Aug-2019. [Online]. Available: https://www.fcc.gov/general/specific-absorption-rate-sar-cellular-telephones. [Accessed: 04-Apr-2020].
[26] A. Kundu, "Specific Absorption Rate evaluation in apple exposed to RF radiation from GSM mobile towers," 2013 IEEE Applied Electromagnetics Conference (AEMC), 2013.
[27] "Dupont Kapton Summery of Properties," Dupont. [Online]. Available: https://www.dupont.com/content/dam/dupont/amer/us/en/products/ei-transformation/documents/DEC-Kapton-summary-of-properties.pdf. [Accessed: 05-Apr-2020].

Performance Analysis of Traditional and SDN Based Handovers in Wireless LAN Networks

Muhammad Emran
Faculty of Informatics and Information Technology
Slovak Technology University
Bratislava, Slovakia
memran@kau.edu.sa

Ivan Kotuliak
Faculty of Informatics and Information Technology
Slovak Technology University
Bratislava, Slovakia
Ivan.Kotuliak@stuba.sk

***Abstract*—With the rapid increase of mobile devices handover management in Wireless Local Area Networks (WLANs) is a concern for user mobility. Traditional handover management methods used for user mobility management are decentralized and do not cater to the delay encountered by the number of Ethernet switches between the Access Points (APs) in the WLAN networks. Software Defined Networking (SDN) technology could fulfill the need of effective centralized handover management and load balancing requirements in the WLAN networks. In this paper, we have analyzed the need for SDN to restructure the current WLAN network architecture. Firstly, we analyzed the existing model and proposed an SDN based model for handover management with improved performance; capable of having multiple switches between the APs during user mobility management. Secondly, we have made an experimental testbed in which multiple switches are placed between the APs connecting the mobile stations. Furthermore, these mobile stations are performing handover during their mobility between the coverage area of the APs on traditional and SDN based WLAN networks. Lastly, the results show the analysis of data rate and delay measured on traditional and SDN based WLAN networks.**

Keywords— SDN, OpenFlow, SDN Controller, OpenFlow Switch, Performance Analysis, WLAN, and Wi-Fi Access Point.

I. INTRODUCTION

WLANs are almost everywhere ranging from homes to enterprise networks and are the need of everyday use. Computers, laptops, tablets, smartphones, smartwatches and home appliances all most every device now support Wi-Fi interface. WLAN architecture is under strain by the use of current Internet applications. Internet applications are resource extensive concerning bandwidth requirement, device configuration, route selection, and troubleshooting. Current WLANs do not meet the user's need as per application requirements; therefore, there is a need for improvements and WLAN architecture should be redefined to meet the demand of the resource-hungry Internet applications [1].

SDN technology is a positive shift towards the separation of the control plane from the data plane of the networking devices. This separation of the control plane from the data plane leads towards the centralized policy management of the SDN based networks. SDN based architecture is capable of managing the WLAN network more efficiently for the need of Internet applications and policy management which will be implemented through the SDN controller using Application Programming Interface (APIs). Mobility management in WLAN networks requires centralized management of the APs, Ethernet switches connecting the APs and the mobile stations (MS) which are in the state of mobility. In current WLAN devices handovers are dependent on the mobile station monitoring and triggering the handover based on the signal strength which is measured in terms of RSSI value. This could not guarantee that the target AP will not be overloaded that might degrade the performance of the communication. The delay in performing the handover is another issue that could also affect the performance of the application running on the mobile device while mobility on using WLAN. So, there is a gap to address these issues and devise some algorithm which could monitor the load on the target AP toward which the mobile device is moving and reduce the delay in handover accordingly so that the performance of the network will be optimized.

In [8] presented model, APs are directly connected with each other without switches involvement. This approach does not seem realistic as the data rate and delay are affected by the presence of the switches between the APs. Existing model does not cater switching delays in its calculation. In this paper, we have proposed an SDN based architecture in which APs are connected using multiple switches between them which are used to provide connectivity of the mobile devices to the WLAN during their mobility. Proposed testbed is developed in Mininet-Wi-Fi which is capable of providing the connectivity of the APs using "n" number of switches and can calculate the delay during the handover of the mobile devices. Based on this model we have connected two APs using 3 switches and measured the data rate and delay on the traditional and SDN based WLAN networks using Iperf and Ping traffic respectively.

The remainder of the paper is organized as follows, Section II, addresses the literature review. In Section III we describe the proposed SDN based handover delay estimation model along with the components used in the architecture. Section IV shows the experimentation setup and implementation of the proposed system. In Section V, results are analyzed and discussed. Section VI concludes the paper with the future work.

II. LITERATURE REVIEW

According to CISCO visual networking index the public Wi-Fi hotspots will grow from 64.2 million in 2015 to 432.5 million by 2020, this figure is sevenfold from 2015 to 2020. In the near future 97% traffic on the Wi-Fi networks will be due to smart devices. Better mobility management will be the need for mobile devices that could lead research in wireless and mobile communication. This rapid growth in the smart devices and increased Wi-Fi traffic has encouraged the researchers in

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

academia to dig into the mobility management standards and come up with better solutions that could handle increasing demand for smart devices mobility management with better features [2].

There are several challenges in current WLAN mobility management standards. Some of them are mentioned here; the association of a MS with an overloaded AP could affect the performance of the MS and could lead to connection lost and throughput degradation. In conventional IEEE 802.11 schemes for wireless APs is the only selection of the MS for the connection establishment with the AP without considering the rest of the parameters for example seamless mobility among adjacent APs and the load on the APs to be selected as target for MS connection establishment [3].

Many researchers are working on the performance enhancement of WLAN by focusing on the improvement in load balancing, association control, signal strength, transmit power and handover. Current enterprise WLAN solutions have centralized control of WLAN infrastructure but these are vendor based proprietary solutions that cannot opt with the changing network requirements. To cope with the changing network conditions, applications and user requirements, its necessary that the WLAN management solutions should be programmable, vendor independent and flexible from a hardware point of view under heavy load conditions and can support seamless handover between WLAN and cellular networks [4].

The delay in the handover is tolerable for browsing or email checking but for real-time applications like video streaming or VOIP sessions it affects the performance. The two reasons for this delay are (i) the client station decides to take handover based on the threshold of either the signal strength which is very low or some frames which are dropped and these threshold values vary based on different vendor's devices which causes variable delays in the detection phase (ii) in the discovery phase when the client station searches for the AP to handover, it scans all the channels of the AP to connect to, will take considerable time as this delay is dependent on the load on AP which is again a vendor-specific and managed issue and needs update patches from the vendor. These issues are the main hindrances in the standardization and do not provide flexibility towards innovations [5].

For video streaming applications an SDN based handover in WLAN is proposed in [6], which uses an optimized scanning procedure to detect the target APs and schedule a handover using OpenFlow protocol to route the traffic from the wired network to the appropriate AP. In [7], it proposes a Centralized and Programmable Wireless Wide Area Network (CAPWAN) protocol to provide seamless handover between WLAN and cellular network. In this solution virtual AP (VAP) is created in the physical AP to abstract the association between the physical AP and the UE. Physical AP is called Soft Access Point (SAP) in which CAPWAN is implemented. VAP holds all the information required for the seamless handover like BSSID, MAC address, UE IP address. The UE architecture is modified in which an agent and an OpenVSwitch (OVS) are created. OVS is having a virtual interface which provides the switching using the bridge between the WLAN and Long-Term Evolution (LTE) physical interfaces. LTE eNodeB has been implemented using the Open-Air Interface open source project.

In [8], it proposed a seamless SDN based WLAN handover architecture which reduces the (i) handover delay, (ii) balances the load on the network by monitoring the loads on the APs and (iii) improves the QoS and resource utilization. The load balancing algorithm selects AP based on 3 factors. The most appropriate AP will be with the strongest signal, a smaller number of connected User Equipment (UE), and have low throughput. In this architecture, propagation delay is reduced by using a pre-transit flow table strategy. In which the flow table of the destination AP is updated by the controller during the handover process to avoid the delay encountered in the Packet-In event by the target AP. But this solution connects the APs directly without having any switch in between the APs. We have enhanced the model presented in [8] by introducing the "n" number of switches in between the APs. According to [8], to reduce or minimize the transmission delay and packet loss during the handover process in vehicular networks, caching technique at base stations is used. In this approach, the SDN-based handover scheme built on a two-level SDN controller architecture which allows the improvement of the network performance.

Inter-slice handover management based on SDN enhances the user's requirements of vehicular communication applications. They are such as latency, the handover call dropping probability, etc. obtained through the inter-slice handover management procedure and architecture given in [10]. In this architecture, the low complexity heuristic algorithm is considered.

A Software-Defined Wireless Networking (SDWN) architecture for IEEE 802.11 networks enhances the performance of the handover in many use cases [11]. The algorithm written for the handover decision process of managing signals allows us to measure the signal-noise ratio between the MS and APs. According to [12], the device-to-device (D2D) handover scheme proved that it can enhance the handover performances compared to the existing scheme. In the SDN-based LTE-A/5G architecture, D2D handover decision method and management maintain the side link stability. A handover algorithm based on D2D was introduced for measuring the handover performance such as data transmission delay.

Paper [13] proposed the 2 types of handover procedures which are global and local for managing LTE and Wi-Fi networks. Although the design of the multi-controller was focused on this handover procedure, the authors used inter handover which is between LTE and Wi-Fi in the local controller area. Here, SDN is not depending on the global controller. According to [14], SDN technology enhanced 3G services and facilities for improving users' requirements. Currently, cellular networks face various limitations. They are such as limited scalability, complex network management, manual network configuration, etc. In the future, 5G based challenges of handover depend on some limitations. They are such as complex and expensive network devices, high cost of emerging technology, the inflexibility of users' requirements, network virtualization, etc.

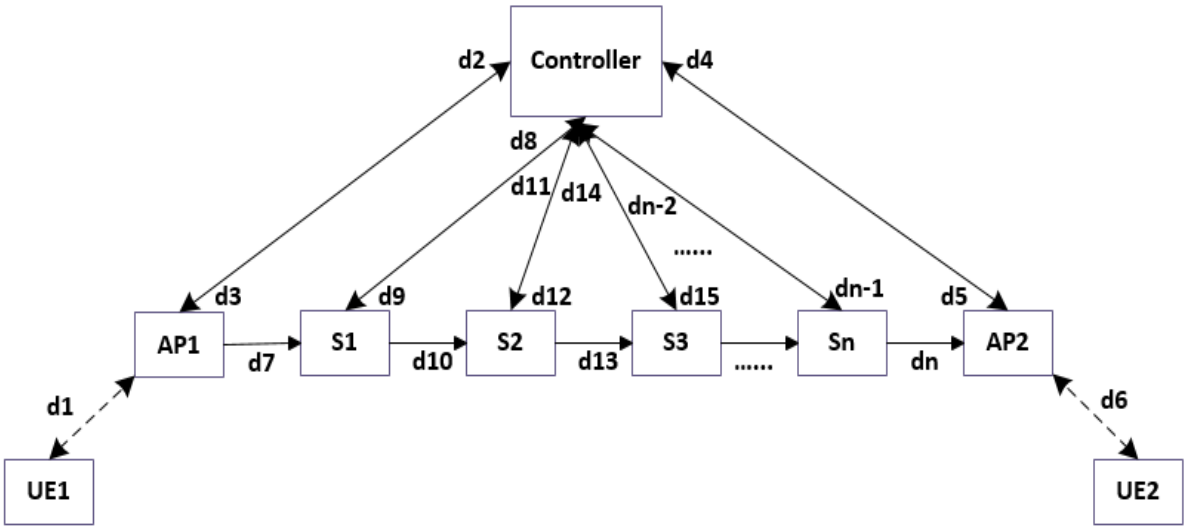


Fig. 1. Block Diagram of Proposed Architecture

III. PROPOSED SDN BASED HANDOVER DELAY ESTIMATION MODEL

To analyze and manage the handover in WLANs an SDN based architecture is proposed in this paper to enhance the performance of the network while performing the handover by estimating the handover delay and managing the load on the WLAN APs. In traditional WLANs, the handover is assisted and managed by the MS whereas in the proposed SDN based WLAN model the handover is performed by the controller which manages and estimates the handover delay in the network while placing multiple switches between the APs. As in some scenarios, it is possible that the APs are not directly connected rather have multiple switches between them to provide the connectivity in multi-stories buildings. Pre transit flow table entry transfer strategy from the controller to respective APs and switches between the APs reduces the delay in the handover of the MSs between the APs in the WLAN.

The proposed system consists of multiple layers, including the Application/Management layer that runs the application which instructs the controller to manage the network. The control layer has an OpenFlow based SDN controller which dictates the flow entries to the OpenFlow based switches/APs. The controller is also responsible to get the feedback from the network devices to further take appropriate actions based on the updates received from the network devices. Infrastructure layer has the data/forwarding plan devices including OpenFlow based switches/APs which receives flow entries from their respective controllers upon inquiring the missing flow information in their flow tables. The descriptions of the components used in the proposed system are given below:

A. SDN Controller:

SDN controller is responsible to gather the information from the network devices and based on this information, manages the network devices to enhance the performance of the network.

B. OpenFlow based AP:

Two APs are used with which two mobile stations are connected. These OpenFlow based APs which are having the OpenFlow tables and are controlled by the Ryu Controller.

C. Mobile Station:

Two MS which are connected to two APs while they are static and later, move and perform a handover.

The delay has a significant impact on the performance of the handover. The block diagram in Fig. 1 shows the delays which will be encountered during the handover process. RSSI, QoS and handover failure rate (Hr) can be calculated using equations 1 to 3 respectively.

Handover is the process of transferring an outgoing call between the evolved nodes (APs) without handover failure. RSSI between UE and the AP is calculated by

$$\mathrm{RSSI} = \mathrm{Tx} - \mathrm{p} * \log(\mathrm{D}) + \mathrm{X} \qquad (1)$$

where, Tx is the transmit power, p is the path loss, D is the distance between the UE and AP and X is the attenuation in dB.

Quality of service depends on the RSSI at the exact location (RSx) of UE, the threshold of the RSSI (RSt), velocity (v), and n is the signal monitoring index. Quality of service will be affected when velocity and its index (Vn) of UE is high and calculated by

$$\mathrm{Vn} = \mathrm{v} * \left(\frac{\mathrm{RSx}}{\mathrm{RSt}}\right)^{\frac{1}{\mathrm{n}}} \qquad (2)$$

Dwell time (T) is the time spent by the UE in the same state. When T > handover time, handover failure will occur. Further $0 < T < 2r/v$, where r is the radius of cell and L is the wired link delay. Using all these calculations and parameters, we can minimize the handover failure rates Hr, which is calculated by

$$Hr = \frac{RSSI * Tx * L}{r * T} \qquad (3)$$

In [8] the delay of the wireless link is $T_w = \frac{D}{C}$, and the delay of the wired link is $T_{nw} = \frac{L}{V}$. When the flow entries in the AP match, the delay T_y is calculated by

$$T_y = \frac{2D}{C} + \frac{L}{V} \quad (4)$$

In the case of the flow table does not match the delay T_z is calculated by

$$T_z = \frac{2D}{C} + \frac{5L}{V} \quad (5)$$

The model presented in the [8] does not cater to the scenario of multiple switches between the APs in the network. In our proposed model we are presenting the solution to calculate the delay T_d in the network having multiple switches between the APs. The proposed model is presented below.

$$T_d = \frac{2D}{C} + \frac{[(n+1) + (4+2n)m]L}{V} \quad (6)$$

where n is the number of switches, m is a variable used to check the flow entries in the flow tables are matched or not. The value of 'm' will be '0' if flow entry is matched and it will be '1' if it does not match. Moreover 'D' is the wireless link delays of UE1 and UE2, 'C' is the data rate of wireless links, 'L' is the wired links delay and 'V' is the data rate on wired links. Assume that the corresponding delays on each link are considered d1, d2, … , dn as shown in Figure 1. d1 & d6 are the delays of the wireless links whereas others are wired link delays.

IV. Experiment Setup and Implementation of the Proposed System

We have used Mininet-wifi emulator for analyzing the performance of the handover of MS among traditional and SDN based WLAN networks as shown in Figures 2 and 3. Positions of MSs (sta1 and sta2) are shown before and after the mobility in Figures 4 and 5. Mobility parameters are kept the same while performing the experiments A and B for measuring the data rate and the delay between sta1 and sta2. The simulation parameters are given in table 1. These parameter values are kept the same in experiments A and B for performing handovers on traditional and SDN based networks.

TABLE I. SIMULATION PARAMETERS

Parameters	Values
Propagation model	LogDistance
Transmission power (dBm)	20
Cell radius (m)	23
Distance between APs (m)	40
Speed of sta1 (m/s)	1.2
Speed of sta2 (m/s)	0.6
Number of switches between APs (n)	3

Experiment A:

In experiment A, the traditional network is setup and two MSs (sta1 and sta2) are connected to AP1 and AP2 respectively as shown in Figure 2. This experiment is conducted without the switches between the APs and later repeated based on our enhanced proposed model, in which the multiple numbers of switches (n = 3) are used between the APs to connect them. In this experiment only the switching delay and wireless propagation delay will be calculated. Three Ethernet switches are providing the connectivity between the AP1 and AP2 which are not shown on these diagrams. APs are kept in the standalone mode to be used without SDN controller in the traditional network topology. Both MSs are in the state of mobility towards the other AP. Sta1 performed the handover from AP1 to AP2 which is assisted by the station itself based on the RSSI values from AP1 and AP2 as shown in Figure 5. Iperf is used to generate the traffic between sta1 on which Iperf server is running and sta2 on which Iperf client is running. The data rate of the generated traffic is measured between sta1 and sta2. This experiment is repeated with the Ping traffic sent and received between sta1 and sta2 to measure the delay.

Experiment B:

In experiment B, the SDN controller is added to experiment A topology as shown in Figure 3. In this experiment the APs are managed by the SDN controller and flow entries are communicated to the APs by the controller. The handover between the APs is being dictated by the controller. Iperf and Ping traffic are generated between the sta1 and sta2 to measure the data rate and delay between the stations. The handover of sta1 from AP1 to AP2 is assisted by the controller.

When the first packet from sta1 arrives at AP1, AP1 misses its flow table entry as AP1 flow table is not having flow entry to forward this packet to sta2. So, AP1 sends this packet to SDN controller in response controller sends the flow entry to the AP1, based on this entry the packet is forwarded and goes through the three switches between AP1 and AP2 and reaches to AP2. Each switch will switch the packet to the controller if it misses its flow table to get the flow entry from the controller. Finally, AP2 receives the packet and also misses its flow table as it is also not having the information to forward this packet to sta2, so AP2 sends this packet to the controller, which in response sends the flow entry to AP2. Based on this flow entry, AP2 forwards this packet to sta2. In this experiment the sum of the delay will include the propagation delays of the APs and the switches to communicate with the controller in addition to the switching and wireless propagation delays. The analysis of the results is discussed in section V.

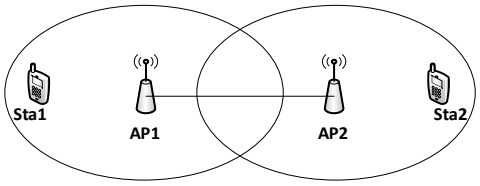


Fig. 2. Traditional Network Topology

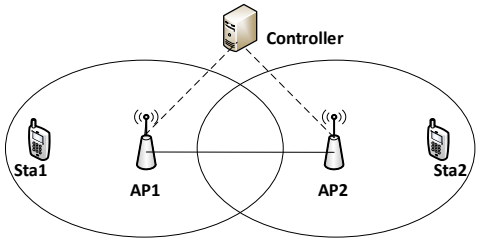


Fig. 3. SDN based Network Topology

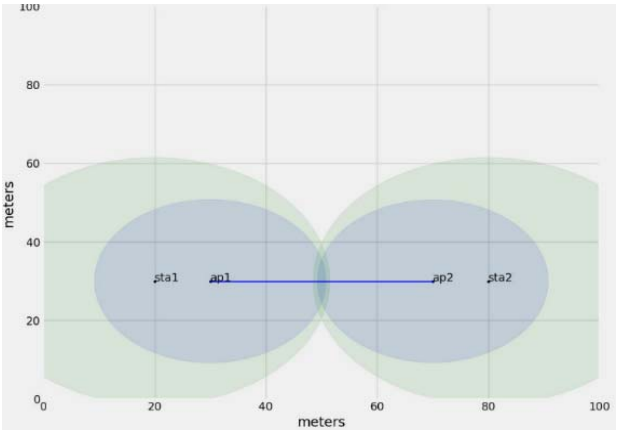


Fig. 4. Stations (sta1 & sta2) before mobility

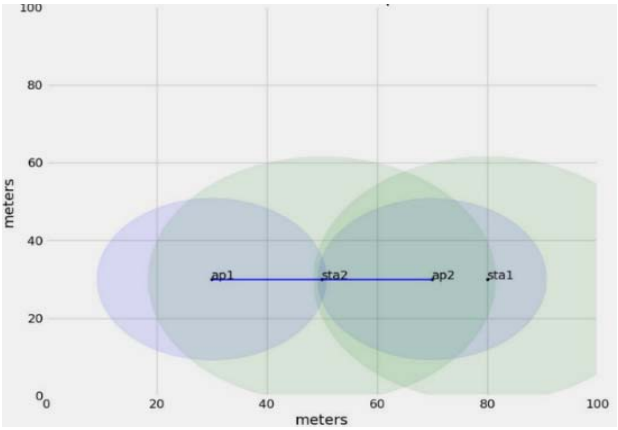


Fig. 5. Stations (sta1 & sta2) after mobility and handover

V. Experiment Results Analysis

Data collected during experiments A and B in section IV for traditional and SDN based WLANs is analyzed and shown below in the graphs. Figure 6 shows the graph for data rate in megabits per second vs time, recorded on traditional and SDN based WLANs and in Figure 7 the graph shows the delay in milliseconds vs Ping packets recorded during the experiments A and B on traditional and SDN based WLANs when the APs are directly connected without having any switches between them. Later based on our proposed model, we repeated the experiments A and B by adding 3 switches between AP1 and AP2 and collected the data. The recorded data is represented on graphs in Figures 8 and 9 showing the data rate and delay recorded using Iperf and Ping traffic.

For clear visibility of the data plotted on the graphs, the logarithmic function is applied to the recorded data. The function used for logarithmic transformation is defined below in equation 7. The value of "α" is selected for the clear prominence of the data shown on these graphs.

$$f(x) = \alpha * log_{10}\, x \qquad (7)$$

where "α" is a multiplicative constant and its value is 0.5 for a graph showing data rate measured in Figures 6 and 8 using Iperf traffic. Figures 7 and 9 show the graphs for measuring the delay using Ping traffic in which the value of "α" is 2.

The data rate of SDN based WLAN is relatively higher than the traditional network as shown in Figure 6 graph. The spikes in the graph show the variation in the values of RSSI which changes due to the mobility of both the MSs (sta1 and sta2).

The propagation model used in these experiments is "LogDistance" which is a built-in propagation model in Mininet-wifi. The fluctuation in the data rate is reasonably minor and depends on the interference in the WLAN during wireless communication.

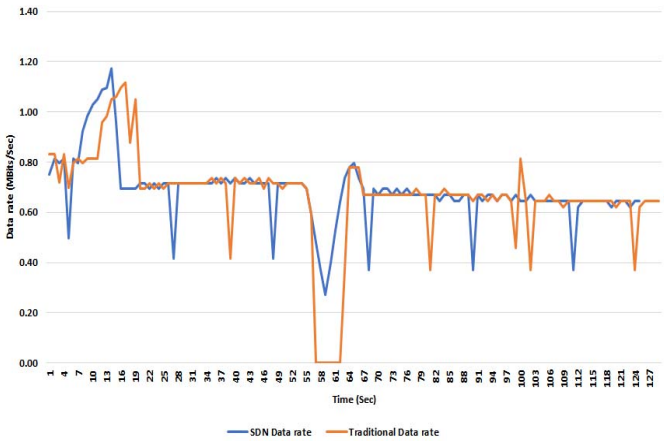


Fig. 6. Data rate vs Time without switches between APs

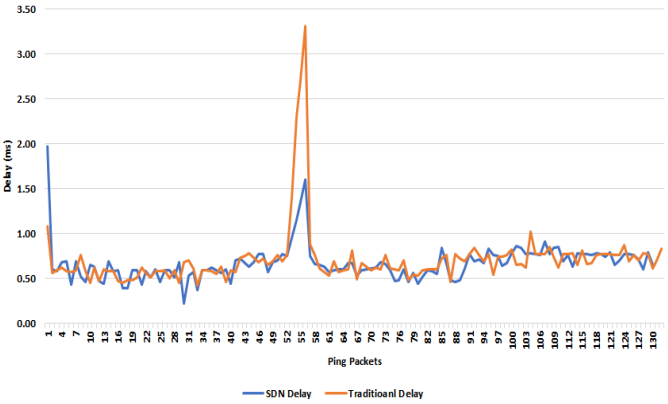


Fig. 7. Delay vs Ping packets without switches between APs

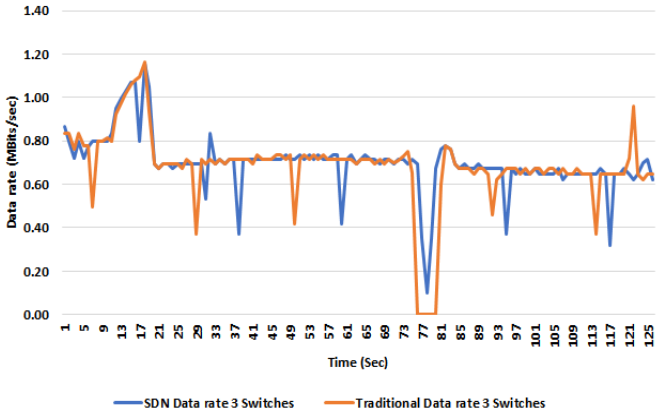


Fig. 8. Data rate vs Time with 3 switches between APs

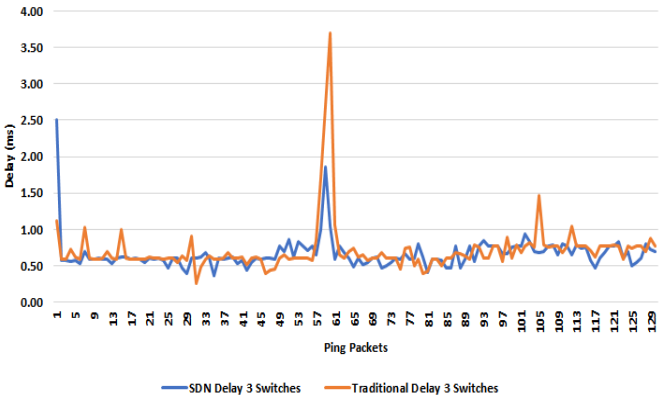


Fig. 9. Delay vs Ping packets with 3 switches between APs

The inherent delay of SDN based WLAN is relatively higher in the beginning as initially flow entries are missed in the flow tables of the APs and the switches between the APs. The initial delay in the SDN based WLAN having switches between the APs is 2.5 ms which is higher than the delay 2.0 ms of the SDN based WLAN without switches between the APs. This clearly indicates that the switching delay incorporates into the overall delay. Later after the handover the traditional network has higher delay as compared to SDN based network which can be seen during the Ping packets 52 to 57 and 57 to 61 in both traditional WLANs without switches and with switches between the AP as plotted on the graphs in Figures 7 and 9 respectively. SDN based WLANs have less delay as compared to traditional WLANs.

VI. CONCLUSION

On the basis of performance analysis of existing model, we proposed an SDN based architecture for seamless handover delay estimation in the WLANs. Our generalized model can be used to calculate the delay (T_d) that occurred in handover during the MS mobility between the APs connected using multiple switches. Our model covers the shortcoming of the existing model in calculating the data rate and delay during the mobility of the MS.

We used the Mininet-wifi emulator to perform these experiments. The experiments are conducted with and without switches between the APs. Our proposed model is generalized for n number of switches between the APs. We performed the experiments connecting the APs directly without switches and repeated the experiments using three switches to connect the APs. While performing these experiments we studied the effect of the switching delays along with the wireless propagation delay and the delay which occurred due to the communication of the APs and switches with the controller when they missed their flow table entries. Based on the analysis of the data gathered during experiments, we concluded that the SDN based WLAN performs better in terms of managing user's mobility and reducing the handover delay as compared to the traditional WLAN despite multiple switches are used between the APs. This indicates that even if the number of switches between the APs increases the performance of the SDN based network will be better than the traditional network in terms of data rate and delay.

In our future work we will evaluate the proposed model by applying it to the LTE network to measure the data rate and handover delay during the mobility of the MSs on traditional and SDN based cellular networks. The discussion about complexity analysis and limitation is out of the scope of the current paper it will be addressed in future work.

REFERENCES

[1] M. Emran, I. Kotuliak, "Real Time Experimentation Study of Software Defined Networking based on Openflow Protocol," in 2nd IEEE Middle East and North Africa COMMunication Conference (MENACOMM), 2019.

[2] Gilani, S. M., Jin, W., Hong, T., Zhao, G., & Xu, C. "SDN-based Handover in Future WLAN", in International Journal of Future Generation Communication and Networking (IJFGCN), Vol.9, No. 12, (2016), pp. 139-154.

[3] T. Hu, K. Xue, W. Wei and W. Jiang, "LENV: A new light-weighted edge network virtualization framework in software-defined wireless networks", in International Conference on Wireless Communications & Signal Processing (WCSP), 2015, pp. 1–6.

[4] Kiran N., C. Yin, Ying Hu, Z. Ali, C. Pan, Israr k., Y. Zhang and G. M., "Handover Based on AP Load in Software Defined Wi-Fi Systems", in Journal of Communications and Networks, Vol. 19, No. 6, (2017).

[5] Jose Q. F, Nailson C, Robertson L, Edisley A, Fernendo M, "A Software Defined Wireless Networking Approach for Managing Handoff in IEEE 802.11 Networks", Hindawi Wireless Communications and Mobile Computing, Vol. 2018, Article ID 9246824, pp. 1-11.

[6] Peter D, Andreas K, Lawrence C, Nicholas B, Nico B, Hans Einsiedler, Christoph P, Daniel M and Miguel S, "A Software-Defined Netorking approach for handover management with real-time video in WLANs", Springer, Vol. 21 (1), 2013, pp. 58-65.

[7] Gang C, Zhaoming L, Tao L, Xiangming W, Luhan W, and Yan Y, "Demo: SDN-based Seamless Handover in WLAN and 3GPP Cellular with CAPWAN", Proc. 13th Int. Symp. Wireless Commun. Syst., 2016, pp. 1-3.

[8] Zhaohui C, Zhaoyang L, Xiaohui D, and Lianming Z, "Terminal handover in software-defined WLANs", EURASIP Journal on Wireless Communications and Networking, Article 68, 2020.

[9] Duo, Ran, Celimuge Wu, Tsutomu Yoshinaga, Jiefang Zhang, and Yusheng Ji. "SDN-based Handover Scheme in Cellular/IEEE 802.11 p Hybrid Vehicular Networks." Sensors 20, no. 4 (2020): 1082.

[10] Mouawad, Nadia, Rola Naja, and Samir Tohme. "Inter-slice handover management in a V2X slicing environment using bargaining games." Wireless Networks (2020): 1-21.

[11] Moura, Henrique, Alisson R. Alves, Jonas RA Borges, Daniel F. Macedo, and Marcos AM Vieira. "Ethanol: A Software-Defined Wireless Networking architecture for IEEE 802.11 networks." Computer Communications 149 (2020): 176-188.

[12] Ouali, Kaouthar, Meriem Kassar, Thi Mai Trang Nguyen, Kaouthar Sethom, and Brigitte Kervella. "An Efficient D2D Handover Management Scheme for SDN-based 5G networks." In 2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC), pp. 1-6. IEEE, 2020.

[13] Park, Junhyuk, and Wonyong Yoon. "SDN-based heterogeneous network architecture with Multi-Controllers." In 2020 22nd International Conference on Advanced Communication Technology (ICACT), pp. 559-561. IEEE, 2020.

[14] Laassiri, Fatima, Mohamed Moughit, and Noureddine Idboufker. "Handover and QoS Parameters a Performance Assessment on 3G Based SDN." In ICT for an Inclusive World, pp. 113-124. Springer, Cham, 2020.

Applicability of Okumura – Hata Model for Wireless Communication Systems in Oman

Mahmoud Al Rashdi
Omantel
Muscat, Oman
alrashdi.mahmood@gmail.com

Zia Nadir
Sultan Qaboos University
Muscat, Oman
*nadir@squ.edu.om

Hassan Al-Lawati
Sultan Qaboos University
Muscat, Oman
hlawati@squ.edu.om

***Abstract*—Current article discusses the usefulness of Okumura-Hata model in urban areas for 2G and 3G communication systems. At some places, applicability of the model is observed and extended to a lower distance from the tower than the specified minimum distance specified in the model. Authors did an extensive drive test and post processed data and proposed theoretical modified model which is presented and compared for different data sets at different locations. 3rd order interpolating polynomial was also used which helps in reasonable prediction for RF engineers.**

***Keywords*—GSM Frequency, Okumura-Hata Model, Pathloss, Propagation models.**

I. Introduction

It is very challenging for wireless network designers to estimate the pathloss of transmitted signals to ensure the quality of service. Generally, the wireless signal is impacted by a variety of propagation mechanisms. These mechanisms include signal reflection, diffraction and scattering. Considering the GSM as the second mobile generation (2G) globally spread RF technology, many propagation models study the pathloss and predict the signal strength at specific distance by predicting the Maximum Permissible Loss (MPL).

This paper proposes and focuses on the urban area with its extension in order to have a suitable propagation model for this area or any other urban area. The Okumura-Hata has been practiced before by several researchers and compared to the real measured data. The Root Mean Square Error (RMSE) is to be used as one indicator as an error indicator and results validator. Okumura-Hata model also needs to be studied for 3G frequency spectrum to provide lower RMSE value which are not done before in these areas. Then, the generalized equation has to be obtained by averaging the RMSE values for studied paths in each area. The generalized formulas than is to be tested and verified in different paths of other areas and the obtained results should be within expected range[1].

For several factors, the RF design is a highly challenging. Special efforts are required in order to provide good coverage with sufficient capacity which satisfies the high demands of this populated area. Hence, the path loss variation with respect to distance is investigated and studied in this work. This is very helpful for RF engineers in order to have a special propagation model that can assist the radio network planners to implement high efficient network.

In a radiocommunication system, the pathloss exponent has a solid effect on the quality of signals. Accordingly, it is required to precisely estimate or predict a perfect design of radiocommunication network. There are several researchers who working in this area. Several research findings are linked or address modelling of pathloss effects for a very narrowband communication by using diverse techniques starting from analytical models to empirical models. Businesses in this area are having a major issue to get maximum received power. So, for them losses that occur during transmission of signals from the transmitter to the receiver is very critical. This work addresses, the empirical method along with other techniques. It is tedious as it involves the huge data collection, analysis and processing and performing drive test. A fixed distance was initially taken from the base station to the receiver and later drive test was conducted.

As a first step, the Centre of Business District (CBD) area has been subdivided into two sub areas. The reason for this is to facilitate the study in terms of data analysis and manipulation. Also, it has to be mentioned that each area differs from each other slightly in terms of buildings pattern arrangement. The buildings of the main area are quite tall and aligned in a uniform pattern. On the other hand, the buildings of the extended area are comparatively shorter and closer to each other. For the purposes of path loss modelling study, a GSM site in each of the two areas has been selected to be used as an RF signal transmitter. The name chosen for the site in the area is 'Site-1' and that for the extension area is 'Site-2'.

II. Propogation Models

A. Free Space Propagation Model

As there is an increase in frequency, further to it, rapidly received signal decreases its amplitude when we increase over distance. Due to this reason, some companies like to use 700Mhz for their cell-phones. This can cover a larger distance. WiMAX service, in the 2.5 GHz band loses power more quickly over distance than 700MHz. Similarly, some services in

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

5800MHz and above lose power even more intensely when distance is increased. This forces the operators to use higher power levels and big high gain antennas which has additional challenges. These figures are not yet in dB which has the effect of reducing the exponential signal deterioration over distance into a straighter graph. The pathloss, which represents signal reduction as a positive quantity measured in dB, is defined as the difference (in dB) between the effective transmitted power and the received power, and may or may not include the effect of the antenna gains. The free space power received by an antenna separated from a transmitter antenna by a distance d, is given by Frii's free space equation (1):

$$P_R(d) = \frac{P_t G_T G_r \lambda^2}{(4\pi)^2 d^2 L} \quad (1)$$

Pt is the conveyed power, Gt , Gr are the Gain of the source and receiver respectively, λ is the wave length(m), d is the distance between source and receiver and L is the system loss factor (L $\geqslant$1). The free-space pathloss denoted by Lp(d) is given by equation (2):

$$L_P(d) = \left(\frac{\lambda}{4\pi d}\right)^2 \Rightarrow L_P(d) = -20 \log_{10}\left(\frac{\lambda}{4\pi d}\right)[dB]$$

$$\because L_P(d) = 20\log_{10} f_C + 20\log_{10} d - 147.56\,[dB] \quad (2)$$

It can be seen from (2) that the free-space pathloss increases by 6dB when doubling the distance.

B. Urban Propagation Environment

There are several parameters which contribute in the propagation mechanisms. The signal is affected as it moves through an urban environment. There are structures, constructions which cause the signals to be returned or diverted and scattered. These phenomena of propagation of radio signals can be caused by Trees and foliage. The reduction of signals strength, can be measured by taking the difference between the median signal levels in front of the building and inside the building. We can go into depth by taking into account the civil engineering aspects of the buildings, materials used etc. In this case we can be more precise in prediction. Reflection, diffraction, and scattering are the three basic propagation mechanisms which impact propagation in a mobile communication system.

C. Plane Earth Propagation Model

Propagation representations are scientific depiction of results of tests carried on the propagation of wave under several different frequencies, antenna heights and locations over different periods and distances. When the radio wave propagates over the ground, it can be partially absorbed and the rest is reflected back to the medium. Due to the reflection from earth surface, the power of the signal can be higher than predicted by the free space model. Pathloss intended for the plane Earth Model is shown in (3 and 4):

$$P_r = P_t\, 20\left(\frac{h_1 h_2}{d^2}\right) \quad (3)$$

$$L_p = -10\ Log\ _{10}\ P_t\, 20\left(\frac{h_1 h_2}{d^2}\right)^2 [dB] \quad (4)$$

Where, (d) = the path length (m)
(h_1) = BS height (m)
(h_2) =MS height (m). We have to select another model due to its limitations for GSM. Model should also consider, all other aspects also e.g. reflections from buildings, multiple propagation or diffraction effects. Further, due to changes in h_2, everything will change and prediction or measurements will not be accurate.

D. Attenuation Factors

Ideally in free space the signal would be sent and received with no loss but that is not the case in reality. As the signal propagates through the medium, it encounters obstacles that contribute in the attenuation of the signal power. Hence, interpreting their contribution in the pathloss equations is necessary. The weather factors are too many and for simplicity they are often neglected. Due to this neglecting, the calculations will encounter some error. To achieve better accuracy, these factors are inserted into the model as RMSE. Rain is one of the attenuation factors. An area with high rain rate would encounter more attenuation than an area with no rain. This attenuation might be expected due to the nature of the rain drop. The rain drop is the result of a condensed water steam at low temperature. Due to its nature, the rain drop might scatter, absorb or reflect the signal. Humidity is another attenuation factor that contributes to the pathloss. The humidity is the result of the evaporation of water in the atmosphere near earth. Areas like coasts are high humidity areas. The attenuation encountered due to humidity in such areas is expected to be higher than areas with low humidity.

III. Empirical Models

Propagation models are elaborated in this section; between them are Okumura-Hata model which was finally adopted for this article.

A. Okumura Model

Okumura's model is one of the most widely used models for signal prediction in urban areas. This model was the result of intensive propagation tests for mobile systems at different frequencies conducted by a Japanese scientist called Okumura. This model is represented by a set of curves (frequency (MHz) verses attenuation (dB)) assuming mobile antenna height of 3m. This model is appropriate under the following circumstances:

- Frequency range of 150-1920MHz,
- Distance from BTS of 1-100km,
- BTS height of 30-100 m.

B. COST231 Model

This model is a modification of Okumura-Hata model. The extension of this model includes higher frequencies which were not covered by the Okumura-Hata model.

C. Walfisch Model

This model was developed by Walfisch and Bertoni. It considers the impact of rooftops and building height by using diffraction to predict average signal strength at street level.

D. Okumura-Hata Model

Hata's model is the extension of Okumura's. This model can be similar to a mathematical illustration of the Okumura's model. It's in the form of empirical formulas. This model is consistent with the Okumura model for distances greater than 1km. The conditions or basic limitations for this are mentioned as below:

- Frequency range of 150-2000MHz,
- Distance from BTS of 1-100km,
- BTS height of 30-200m,
- MS antenna height 1-10m

The limits on this technique is due to range of test outcomes as illustrated above [1]. Hata created several typical Pathloss scientific models for different areas e.g. urban, suburban and open country environments, as mentioned in the following equations, respectively. Model takes urban areas as a reference and applies correction factors as mentioned below:
For urban areas generalized Okumara-Hata model is given below:

$$L_{dB} = A + BLog_{10}R - E_{1,2,3} \quad (5)$$

Where:

$$A = 69.55 + 26.16\,Log_{10}f_c - 13.82\,Log_{10}h_b$$

$$B = 44.9 - 6.55\,Log_{10}h_b$$

$$E_1 = 3.2\left(Log_{10}\left(11.7554h_m\right)\right)^2 - 4.97 \quad for\ cities;\ f_c \geq 300MHz,$$

$$E_2 = 8.29\left(Log_{10}\left(1.544h_m\right)\right)^2 - 1.1 \quad for\ cities;\ f_c \leq 300MHz,$$

$$E_3 = \left(1.1Log_{10}f_c - 0.7\right)h_m - \left(1.56\,Log_{10}f_c - 0.8\right)$$

for relatively smaller cities.

Where:
h_m; MS height [m]
d_m; distance between the phone and the building [km]
h_0; height of a building [m]
h_b; BS height [m]
r; circle distance BS and mobile [m]
R=r x 10^{-3} great circle distance between BS and mobile [km]
f_c= frequency [MHz]
Rest of the parameters are as mentioned above for small and medium cities[2-4].

E. Extended Sakagami model

This model is described in [5] is presented by following equation:

$$\begin{aligned} L_p[dB] = {} & 101.0 - 7.11\log_{10} W \\ & + 7.5\log_{10} H \\ & - \left(24.37 - 3.7\left({}^{H}/_{h_b}\right)^2\right)\log_{10} h_b + (43.42 \\ & - 3.11\log_{10} h_b)\ \log_{10} d \\ & + 20\log_{10} f_c - a(h_m) \end{aligned}$$

Where $a(h_m)$: correction factor for antenna height (h_m).

$$a(h_m)[dB] = 3.2(\log_{10}(11.75h_m))^2 - 4.97$$

So Pathloss can be calculated using the following equation:

$$L_p(dB) = P_t - P_r \quad (6)$$

Where P_t is the transmitted power which is equal to 47dB for 2G and 34dB for 3G transmitters and P_r is the received power. The major models used for this study is Okumura-Hata model. Authors contribution is the usage of the model for less than 1km and proposing an interpolated polynomial for RF planners.

IV. Results and Discussions

An intensive drive test was conducted along all pre-identified paths using TEMS. Furthermore, the positioning information is collected via a GPS antenna. The measured data for each path has been recorded in terms of log files and processed using ACTIX. Google Earth program (not shown here in this paper) was also used to plot the received signal strengths. Duplicate data sets were also cleaned before post processing. After that, the data has been worked upon. This area of CBD can be considered as an urban area. After defining, the study was supported to make an evaluation among the tentative and hypothetical data and the outcome is as shown in following Figures. 1-4:

The results of measured path loss have been used as an input for usage with developed MATLAB scripts in order to plot the measured path loss and the predicted path loss curves as a function of distance. MATLAB is an efficient software for mathematical calculation and data analysis. After that, the variation between measured and predicted results has been obtained using RMSE method which was introduced earlier. The RMSE error indication technique has been used to modify each data set to have improved prediction capabilities. RMSE value combined all known and unknown factors and parameters impacting the path loss amount of the propagated signal. These factors include multipath effects, various propagation mechanisms and different weather conditions e.g., temperature and humidity. The RMSE value has been utilized to establish the modified propagation model for each studied path. After that, the RMSE value has been recalculated for the modified equations.

<u>*Data-Set-A*</u>
f_c=935.4MHz; h_b=25m; h_m=1.5m; H=30; W=20;

RMSE Value for Okumara-Hate Model was 7.3864dB and modified RMSE value after modification of Okumara-Hata model was 5.1321dB which is within acceptable range [6]. Following figures (1-4) details the Pathloss and modeling.

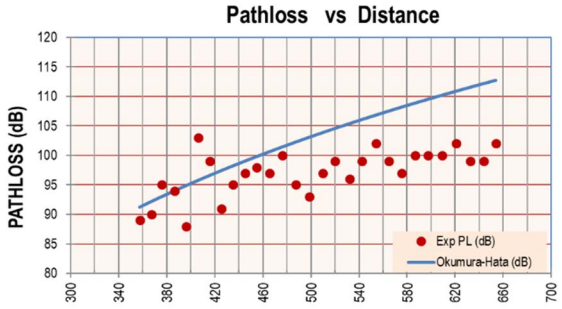


Fig. 1. Experimental and Theoretical Pathloss

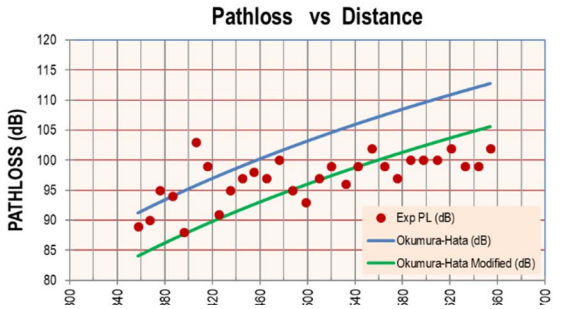


Fig. 2: Exp. And Theoretical Pathloss with modified model

Data-set-B

f_c=2130MHz; h_b=28m; h_m=1.5m; H=30; W=20;

RMSE Value for Okumara-Hate Model was 15.6096dB and modified RMSE value after modification of Okumara-Hata model was 4.05838dB which is within acceptable range.

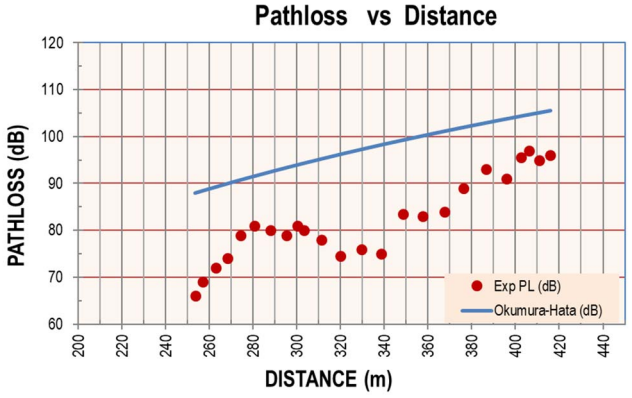


Fig. 3: Theoretical and Experimental Pathloss versus distance

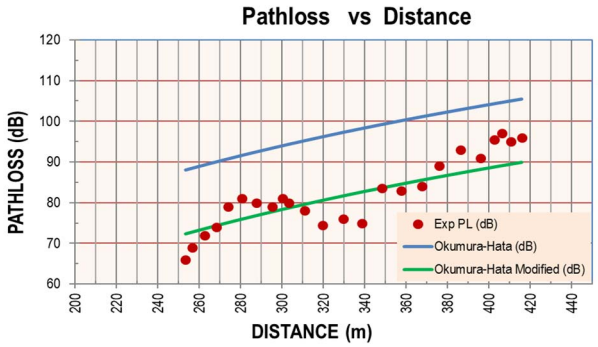


Fig. 4: Theoretical and Exp. Pathloss versus distance with modified model.

Above shown graphs and after comparing with articles [2, 3] the results visibly display that the measured Pathloss is smaller in value than the predicted Pathloss by a variance from 5dB to 20dB. Nevertheless, there are several explanations that may have triggered those substantial modifications. First of all, in Japan there are few zones virtually fulfilling the conditions; and if any, they are narrow. Because of that reason Okumura selected the value for urban area as standard for open areas [7]. Furthermore, the topographical situation of Japan is different from that of our country due to geographical differences. Accordingly, Root Mean Square Error (RMSE) was considered amongst those two different values of pathloss, for Hata model, using following (7) [1, 6]:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Pmeasured_i - Pcalc_i)^2}{(N-1)}} \quad (7)$$

N: Measured Data Points

The RMSE acceptable range is up to 6dB so, the RMSE is adjusted with the Hata equation for urban area and the modified Hatas' equation is as given below (8):

$$\begin{aligned} L_{p\_mod}(urban) &= 69.55 + 26.16Log_{10}(f) \\ &\quad -13.82Log_{10}(h_b) + (44.9 - 6.55Log_{10}(h_b))Log_{10}(d) \\ &\quad \pm MSE - (1.1Log_{10}(f) - 0.7)h_m - (1.56Log_{10}(f) - 0.8) \end{aligned} \quad (8)$$

The modified result of Hata equation is shown in Fig. 5 and Fig. 6 and the RMSE in this case is less than 6dB, which is acceptable[6].

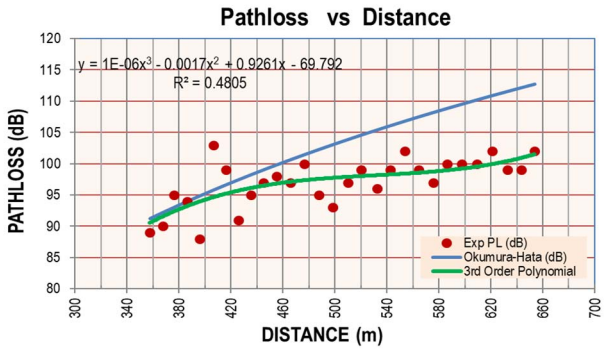


Fig. 5: Theoretical and Experimental Pathloss versus distance for data-set-A

For verification, whether the modified Hata's equation is applicable for some other areas, another data generated from TEMS tool has been used. Based on that practical data, the propagation Pathloss and the distance have been re-verified for [4].

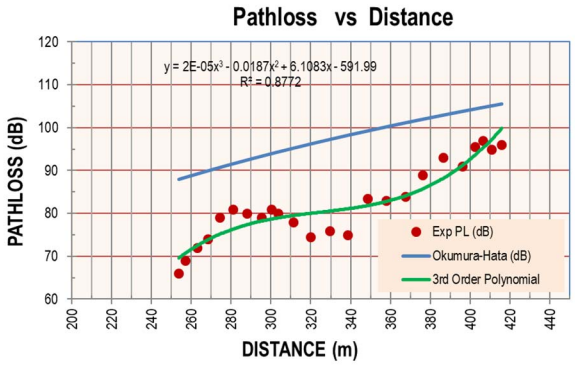


Fig. 6: Theoretical and Experimental Pathloss versus distance for data-set-B

Hypothetical imitation and the found new data are equated and examined further using 3rd order polynomial to interpolate on same experimental data set which provided moderate flexibility. As shown in figures above, higher order polynomial also provides an equally even and effective guess without cumulative computational difficulties, which is expected for these estimates. A decent relationship is detected for the entire series of data set. The decent agreement of the features demonstrate that experimentally replicated numbers is a decent representation of that defined by Hata model.

Additionally, the simulation and the attained tentative data is associated and examined additionally using other models like extended Sakagami model on the set of the new data which gives also satisfactory results (Not the scope of the current paper). After observation, it can be safely said that the scatter plot of the experimental data on pathloss vs distance reveals a third order polynomial tendency. Fig.5 and Fig. 6 above, show the theoretical, experimental and 3rd order polynomial plots for Okumura-Hata propagation model. As can be seen, results show good agreement between various studies.

Universally, by calculating the RMSE for the second dataset it was also found to be less than 6dB, a satisfactory number. Nevertheless, some experimental values were a bit far from the interpolated points that can be related to the nature of the cell with high rise structures. Though there are many forecasts approaches that are based on deterministic procedures through the availability of improved databases, the Okumura-Hata model is still frequently used [8] [9]. We are also aware that the hindrances in the path significantly affect the radio signal propagation [10]. Above all, wireless communication system avoids obstacles such as crossing objects owned by others. There are also many difficulties in establishing a wireless communication system in some applications [11]. As can be seen from above results, there are some discrepancies between RMSE values for each path compared to the other, although the modified results for each path are within accepted standards. After deep analysis and investigation of the possible reason for this mismatch, it has been noticed that some paths are served by the main lobe of antenna radiation, some are served by side lobes and some are served by a combination of the main lobe and side or back lobes. This conclusion has been obtained by reviewing all paths in Google Earth map with respect to the area of specific serving lobe. The following figure 7 shows general antenna radiation pattern concept and the radiation pattern of another site.

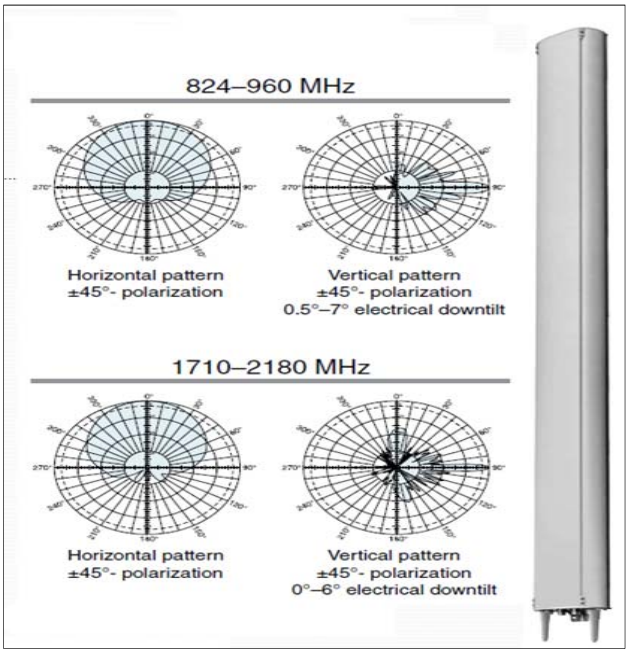


Fig. 7: Radiation pattern of an antenna on one site

In order to produce a generalized model for UMTS path loss prediction use, the average of RMSE for all tested paths was calculated for the model. Then, the same well known process has been applied in to obtain a generalized equation for the propagation model. After that, these equations have been applied for three different paths but not shown here. They are namely: Site-1 Path-3, Site-2 Path-2 and Site-3 respectively. The MATLAB programs for each path was developed and used for generalization and further verification.

This work can be considered as a step forward in establishing generalized propagation models used for path loss prediction in other urban cities of Oman. Further intensive and comprehensive studies and research is recommended to achieve this goal. Also, it is highly recommended to incorporate various modelling techniques in addition to the RMSE method to study their accuracy and impact. Moreover, it is recommended to include more parameters in future prediction models that impact the signal propagation e.g., antenna parameters and patterns.

V. Conclusion

This study focused on forecasting the root mean signal strength in diverse areas. As utmost propagation models aim to forecast the median pathloss, existing prediction models vary in terms of their applicability over different terrain and environmental conditions. The effects of terrain situation predicted at 900MHz and 2.1GHz were analyzed. Experimental outcomes of radio signals propagation for an urban area in Oman were related with those predicted based on Okumura-Hata model. The contribution is the prediction by at lower distances than the model is generally used and also validation of experimental data. If precise environmental information was included in the model, better prediction results might be achieved. 3rd order polynomial gave us also the unavailable experimental points showing a good agreement within adequate boundaries.

Acknowledgment

The authors would like to thank SQU and Omantel for their support for this study.

References

[1] Shalangwa, S. D. A., and S. K., "Measurement and modeling of pathloss for GSM900 in sub urban environment over Irregular terrain," *IJCSNS International Journal of Computer Science and Network Security,* 2010.

[2] Z. Nadir, "Empirical pathloss characterization for Oman," in *2012 Computing, Communications and Applications Conference*, 2012, pp. 133-137.

[3] Z. Nadir and M. Ahmad, "Characterization of Pathloss Using Okumura□Hata Model and Missing Data Prediction for Oman," *AIP Conference Proceedings,* vol. 1285, pp. 509-518, 10/15 2010.

[4] Z. Nadir, N. Mohamed, and F. Touati, "Pathloss Determination using Okumura-Hata Model and Spline Interpolation for missing Data for Oman," *Lecture Notes in Engineering and Computer Science,* vol. 2170, 07/01 2008.

[5] T. Fujii, Y. Ohta, and H. Omote, "Empirical Time-Spatial Propagation Model in Outdoor NLOS Environments for Wideband Mobile Communication Systems," in *VTC Spring 2009 - IEEE 69th Vehicular Technology Conference*, 2009, pp. 1-5.

[6] W. Jianhui and Y. Dongfeng, "Propagation measurements and modeling in Jinan city," in *Ninth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (Cat. No.98TH8361)*, 1998, pp. 1157-1160 vol.3.

[7] R. D. Wilson and R. A. Scholtz, "Comparison of CDMA and modulation schemes for UWB radio in a multipath environment," in *GLOBECOM '03. IEEE Global Telecommunications Conference (IEEE Cat. No.03CH37489)*, 2003, pp. 754-758 Vol.2.

[8] Z. Nadir and M. Bait-Suwailam, "Pathloss Analysis at 900 MHz for Outdoor Environment," presented at the International Conference on Communications, Signal Processing and Computers (EUROPMENT 2014), Interlaken, Switzerland, , 2014.

[9] Z. Nadir, M. Bait-Suwailam, and M. Shafiq, "RF Coverage Analysis and Validation of Cellular Mobile Data using Neural Network," *INTERNATIONAL JOURNAL of NEURAL NETWORKS and ADVANCED APPLICATIONS,* vol. 1, pp. 30-36, 01/01 2014.

[10] A. Hrovat, G. Kandus, and T. Javornik, "Path loss analyses in tunnels and underground corridors," *International journal of communications,* vol. 6, pp. 136-144, 08/01 2012.

[11] L khagvatseren. T and H. F, "Path loss aspects of a wireless communication system for sensors," *International Journal of Computers and Communications (NAUN),* vol. 5, 2011.

Double Mobility WSN: Exploiting the Mobility of Sink and Cluster Head Nodes for Better WSN Energy Preservation and Lifetime

Ala' F. Khalifeh[1] , Husam Abid[1] , Khalid A. Darabkh[2]
[1]Dept. of Electrical and Communication Engineering, German Jordanian University, Amman, 11180, Jordan
Email: (ala.khalifeh@gju.edu.jo, h.abed266@gmail.com)
[2]Dept. of Computer Engineering, The University of Jordan, Amman, 11942, Jordan, Email: k.darabkeh@ju.edu.jo

***Abstract*— Mobile Sensor Nodes (MSNs) have been recently proposed to leverage the performance of conventional static Wireless Sensor Network (WSN). In this paper, we propose to energize both the sink and Cluster Head nodes with mobility, such that wireless communication path loss is minimized thus preserving the sensor nodes energy and extending the network lifetime. The performance evaluation of the proposed scheme shows a superior performance over other related work in terms of extending the network lifetime, and preserving the nodes energy.**

***Keywords*— WSNs, CHs, mobile sensor node, mobile sink.**

I. INTRODUCTION

Wireless Sensor Networks (WSNs) have been flourishing in the past few years especially with the advent of many Internet of Things (IoT) based applications and scenarios [1]. This is due to the fact that WSNs can play a vital role in providing the needed infrastructure for scaling up the connectivity, availability and security of these IoT- based applications. For instance, a smart meter can be connected to a WSN to send its sensed data to a remote sink node, where data is analyzed and logged. This communication scheme can be done without utilizing the telecom infrastructure, where a WSN can be a cheap and effective substitute, especially that WSN has the capability of sending the sensed data for long distances utilizing a multi-hop mesh network, thus having scalable and cost-effective solution. Another potential application for WSN is to sense and send data about an Area of Interests (AoI) for further analysis and decision-making. For example, WSN can be deployed in an agricultural farm, where sensor nodes can be distributed to collect weather, soil related data, and send it for a sink node for further processing and analysis. Conventional WSN consists of static Sensor Nodes (SNs) distributed over the AoI, each node is equipped with sensors, a microcontroller, communication module and battery. Given the fact that these nodes are often positioned in hard to reach places, charging the nodes' batteries may not be feasible, therefore, optimizing the energy consumption of the sensor nodes is paramount to their operation [2, 3]. Several techniques are used to preserve the precious energy source of these nodes, of which clustering is a widely used and effective solution [4]. SNs are grouped together forming a cluster, each cluster has a Cluster Head (CH) that has a superior power source, and may have extra communication modules and interfaces, and thus it is considered as a power node. The CH collects the sensed data from the SNs, processes them and decides whether the packets need to be sent to the remote sink node for further processing and analysis. Further, the CH node can compress the packets to be sent for the sink node for further power optimization.

In the literature, the integration of the Mobile Sink Node (MSN) [5] to the WSN is performed in order to decrease the network energy consumption and reduce the number of transmission hops needed to reach the sink node. In particular, the MSN moves to each CHs, and when the distance between them approaches a threshold (usually a small distance), the CH sends the collected data to the it. In this scheme, the MSN adapts its transmission energy proportionally with respect to the threshold distance, thus resulting in significant conservation of the communication energy, especially when compared with the energy consumed due to the long established communication link with the stationary remote sink node.

Based on this concept idea, the authors [6] developed a novel optimized clustering framework to investigate the effect of sink mobility, where two routing protocols are proposed based on the spiral mobility patterns. The first one is the Spiral Mobility based on Optimized Clustering (SMOC) for optimal data extraction for a single mobile sink, and the second one is the Multiple Sink Based on SMOC (M-SMOC) for large-scale WSNs. In SMOC, a single mobile sink moves in a spiral pattern over the sensing field to collect data from sensor nodes and CHs, while in M-SMOC, several mobile sinks (one mobile sink node for each cluster) are utilized for data collection, to reduce the delay while collecting the sensors data.

The authors in [7] introduced a novel data gathering method based on utilizing both clustering and mobile sink. The proposed method uses clustering concept to get energy efficiency and a mobile sink to visit every cluster and collect sensed data of the SNs. After visiting all the clusters, it returns to the base station where data is uploaded and further processed. Form the conducted literature work; it is noticeable that most of the proposed algorithms utilizes a mobile sink node to reduce the CH communication energy consumption.

In this work, we propose a new WSN architecture called Double Mobility- WSN (DM-WSN), where not only the sink node is mobile, but also the CHs, where the CHs moves within their clusters in order to find the optimal location with respect to the cluster SNs, which will result in a reduced energy consumption. Indeed, this paper continues our previous work that aims at building a secure, energy efficient and scalable WSN [8-10]. In particular, this paper extends our recently

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

published papers [11, 12] where the CHs optimal locations are derived and computed such that the new CHs positions minimize the path-loss between the CH and the sensor cluster nodes. Further, the spiral pattern for the CHs positions allocation is adopted due to its energy efficiency [13], where the CHs are initially deployed in predefined locations in a spiral pattern within the AoI.

The rest of the paper is organized as follows: Section II describes the proposed double mobility WSN architecture. Section III depicts the performance evaluation of the proposed architecture. Finally, section IV concludes the paper and proposes future work.

II. The Double Mobility WSN architecture

The proposed DM-WSN architecture is depicted in Figure 1. In this network, the SNs go through three different stages as follows: the nodes' deployment and clustering, CH location optimization and re-allocation, then the mobile sink node navigation and data collection.

A. Stage 1: Nodes' deployment and clustering

In this stage, the SNs and CHs are deployed in the AoI. First, the optimal number of CH nodes (K_{opt}) is calculated using Eq. (1) [14], where n is the total number of SNs in the network.

$$K_{opt} = \sqrt{\frac{n}{2\pi}} \left(\frac{2}{0.765}\right) \quad (1)$$

Calculating K_{opt} estimates how many clusters is needed to cover the AoI. Then, the locations of CHs are distributed over a spiral trajectory, which reported to guarantee an optimal nodes' coverage for the AoI while balancing the nodes energy consumption, which in turn will extend the network lifetime [13, 15]. The CH can be deployed using an Unmanned Arial Vehicle (UAV) or airplane as presented in [16]. The nodes' deployment process is described as follows: If the distance between any two CHs is equal to $\sqrt{3}\,Rs$, where Rs is the CH coverage radius, then there should be no hole between the adjacent CHs thus guaranteeing minimal overlap in the area. The process of CH nodes' deployment is repeated for every $\sqrt{3}\,Rs$, over the spiral trajectory until covering the entire AoI. In Figure 2, eight CHs are deployed (CH1, …, CH8), the figure also depicts the SNs for CH4, CH5 colored in blue, yellow, respectively. Further, the initial and optimal locations of CH5 is shown using green, red color circles, respectively. Therefore, the CH will move from its initial to the optimal location.

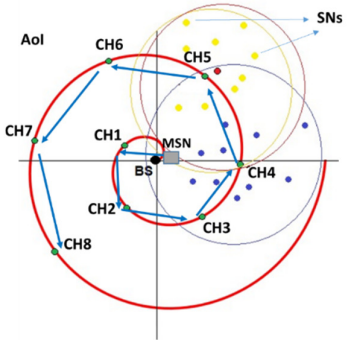


Fig. 2 The Double Mobility WSN Architecture

After deploying the CH in their pre-defined locations over the spiral trajectory, the SNs form several clusters by associating themselves with the CHs that have the strongest Received Signal Strength Indicator (RSSI) beacon signals.

B. Stage 2: CH location optimization and re-allocation

In order to reduce the nodes' energy consumption while communicating with the CH, an optimization problem for finding the optimal location of the CH with respect to the SNs has been formulated and solved in our recent work [11, 12]. The main objective is to find the optimal CH location that minimizes the path loss between the SNs and the CH within the cluster. The optimal location of the CH i (x_i, y_i) is calculated using Equations (2, 3).

$$x_i = \frac{\sum_{k=1}^{Ni}(x_k)}{Ni-1} \quad (2)$$

$$y_i = \frac{\sum_{k=1}^{Ni}(y_k)}{Ni-1} \quad (3)$$

Where Ni is the number of nodes of cluster i, and x_k, y_k are the coordinates of the cluster nodes. Once the CH moves to its optimal location, some SNs may become out of CH coverage, thus after the CH motion, it discovers the out of reach nodes, and both the CH and SNs adjust their transmission energy to retain the connectivity. The process of CH repositioning is continuously repeated during the network lifetime once the network topology is changed, which happens due to the energy depletion of some SNs.

C. Stage 3: Mobile sink nodes navigation

Once the clusters are formulated, the SNs start sending the sensed data to the CHs, where captured data is stored until the distance between the mobile sink node location and the CHs less than or equal a certain threshold, which should be small to reduce the transmission energy between the CH and the MSN. Further, while the MSN offloads the captured data from the CH, the SNs should be able to continue sending the captured data to the CHs. As such, the CH should have two radio communication modules, each work on different frequency band, thus allowing simultaneous transmission of the SNs to the CHs and from the CHs to the MSN. As depicted in Fig. 2, the MSN leaves the BS and moves to CH1, the MSN movement speed adapts itself such that the trip time (T_t) between any two adjacent CHs are fixed to a certain duration, which is equal to the SNs Cluster Round Time (CRT). The CRT is defined as the time needed to allow all SNs within the cluster to send their captured data once to the CH utilizing Time Division Multiple Access (TDMA) scheme. Once the CRT is defined, the allocated time slot (t_{si}) for each SNs of the cluster i can be calculated using Equation (4).

$$t_{si} = CRT / N_i \quad (4)$$

III. PERFORMANCE ANALYSIS

To assess the performance of the proposed MSN and CH mobility schemes, a simulator for WSN is built using Matlab. In what follows, the simulation setup, parameters and metric used to evaluate the effectiveness of the proposed scheme is summarized, followed by the simulation and performance evaluations results.

A. *Simulation setup*

Table 1 shows the main parameters used in the simulation. Further, the performance of the Double Mobility WSN (DM-WSN) scheme is compared with our previous work (the Multi-Cluster Head Positioning Algorithm (Multi-CHP) [12], where the CH changes its location optimally with respect to the cluster nodes, without having a mobile sink node. Further, the DM-WSN is compared with the Low Energy Adaptive Clustering Hierarchy (LEACH) [16], which is one of most earliest clustering protocols algorithms that utilizes dynamic random rotations for the CHs for each round. LEACH works toward balancing the energy consumption over the network thus the nodes' energy consumption is improved.

Table 1: The simulation parameters

Description	Value
Number of nodes	600
Area	100 × 100 m2
Transmitter electronics (ETX-elec) Receiver electronics (ERX-elec) (ETX-elec) = (ERX-elec) = (Eelec)	50 nJ/bit
Energy consumed by the amplifier to transmit at a shorter distance ϵ_{fs}	10 pJ/bit/m^2
Energy consumed by the amplifier to transmit at a longer distance ε_{amp}	0.0013 pJ/bit/m^4
Total initial energy	350 J
Sink node location	(0, 0)
Packet size	2000 bits

In the conducted simulations, the total number of nodes n equals to 600, which resulted in the following number of CHs:

- For LEACH, the number of CHs = 5% × 600 = 30.
- For the spiral CHs distribution, the number of CHs equals to 26 according to Equation (1).

Further, in order to ensure a fair comparison with all other related works, the total initial energy for all nodes of the multi-CHP and DM-WSN are set to be equal to the total initial energy of the SNs in LEACH protocol. The initial energy was set to 350 J, which is distributed as follows:

- Each SN in the LEACH protocol has an initial energy equals to 0.583 J.
- Each SN in multi-CHP and DM-WSN has an initial energy equals to 0.5 J.
- Each CH has an initial energy equals to 2.432 J.

B. *Evaluation metric and results*

Three main metrics were used in the comparison. The total consumed energy, the total residual energy, and finally, the total number of dead nodes, all as a function of the number of rounds. As shown in Figs. 3-5. The proposed DM-WSN significantly outperforms the multi-CHP and LEACH protocols. Furthermore, three other well-known comparison parameters widely used in the context of WSN are depicted, which are derived from Figure 5. Particularly, they are the First Node to Die (FND), Half Node to Die (HND), and Last Node to Die (LND) performance metrics.

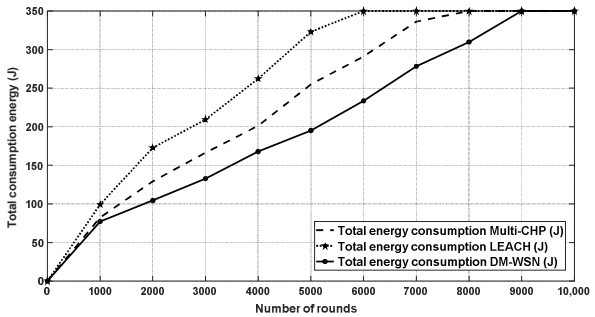


Fig. 3: Energy consumption for the proposed algorithm and other related work.

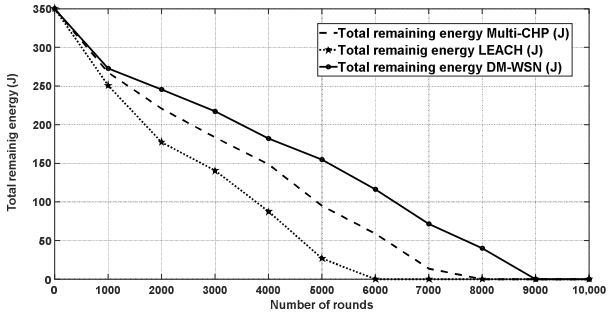


Fig. 4: Residual energy for the proposed algorithm and other related work.

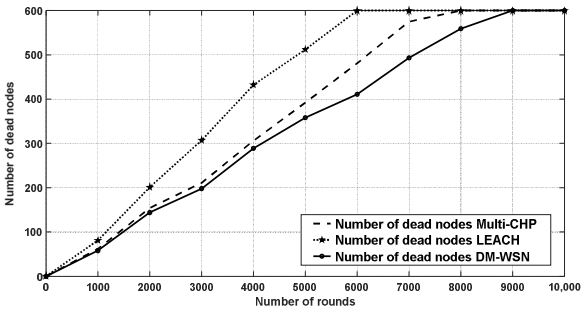


Fig. 5: Number of dead nodes for the proposed algorithm and other related work.

Table 2 shows how the proposed DM-WSN outperforms the multi-CHP and LEACH protocols in terms of both HND and LND. However, for the FND metric, the

DM-WSN was slightly outperformed by the multi-CHP but it largely outperforms the LEACH protocol. The superb performance of the proposed DM-WSN proves that proposed DM-WSN conserved the nodes' energy for larger number of rounds when compared with the other two protocols.

Table 2: The FND, HND, and LND metrics for the different scenarios of Figure 5.

Scenario/Metric	FND	HND	LND
multi-CHP	80	4000	8000
DM-WSN	76	4160	9000
LEACH	55	3200	6000

IV. CONCLUSION AND FUTURE WORK

This paper presented a new wireless sensor network architecture that leverages the mobility of both the cluster head and sink node in order to reduce the communication energy consumed by the sensor nodes and cluster heads. This is achieved by first moving the cluster head node to the optimal location with respect to the cluster nodes, such that the communication path loss is minimized. Second, the sink node is proposed to be a mobile node that travels to each cluster and collects the captured sensor nodes data, which is stored at the cluster head. The close proximity of the CH and mobile sink node reduces the amount of energy needed to transfer the captured packets from the CH to the sink node, thus significantly enhancing the WSN lifetime. As a future work, we are working toward implementing the proposed Double Mobility WSN and conducting real-experiments to compare it with the ones obtained by simulations.

ACKNOWLEDGMENT

This work is supported in by the North Atlantic Treaty Organization (NATO) SPS Project No. SPS G4936 and by the German Jordanian University under the grant No. SEEIT 02/2018.

REFERENCES

[1] S. Bansal, D. Kumar, "IoT Ecosystem: A Survey on Devices, Gateways, Operating Systems, Middleware and Communication", International Journal of Wireless Information Networks, 2020, 1-25.

[2] Y. Jaradat, M. Mohammad, I. Jannoud, and D. Zaidan, "The Impact of Nodes Distribution on Energy Consumption in WSN", in 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), pp. 590-595. IEEE, 2019.

[3] M. Al-Shalabi, M. Anbar, T. Wan, and A. Khasawneh, "Variants of the Low-Energy Adaptive Clustering Hierarchy Protocol: Survey, Issues and Challenges". Electronics, 2018, 7(8), 136.

[4] M. Masoud, Y. Jaradat, D. Zaidan, and I. Jannoud, "To Cluster or Not to Cluster: A Hybrid Clustering Protocol for WSN", in 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT) (pp. 678-682).

[5] D. Licea, M. Ghogho, D. McLernon, and S. Zaidi, "Mobility Diversity-Assisted Wireless Communication for Mobile Robots", IEEE Transactions on Robotics, 32(1), 214-229..

[6] M. Asad, Y. Nianmin, M. Aslam, "Spiral Mobility Based on Optimized Clustering for Optimal Data Extraction in WSNs", Technologies 2018, 6, 35.

[7] D. Mamurjon, B. Ahn. "A Novel Data Gathering Method for Large Wireless Sensor Networks", IERI Procedia, 4, 288-294.

[8] A. Khalifeh, N. Bartolini, S. Silvestri, G. Bongiovanni, A. Al-Assaf, R. Alwardat, and S. Alhaj-Ali, "Hybrid Wireless Sensor Networks: A Prototype", In IFIP Conference on Human-Computer Interaction, pp. 549-553, 2019.

[9] A. Khasawneh, O. Kaiwartya, A. Khalifeh, L. Abualigah and J. Lloret, "Green Computing in Underwater Wireless Sensor Networks Pressure Centric Energy Modeling", In IEEE Systems Journal, doi: 10.1109/JSYST.2020.2996421.

[10] A. Khalifeh, K. Rajendiran, K. Darabkh, A. Khasawneh, O. AlMomani, Z. Zinonos, "On the Potential of Fuzzy Logic for Solving the Challenges of Cooperative Multi-Robotic Wireless Sensor Networks", Electronics 2019, 8, 1513.

[11] A. Khalifeh, H. Abid, K. Darabkh, "Improving Energy Conservation Level in WSNs by Modifying CH Node Location", in the proceedings of the Sixth International Workshop on Internet of Things: Networking Applications and Technologies (IoTNAT 2020), Paris, France, 20–23 April 2020.

[12] A. Khalifeh, H. Abid, K. Darabkh, "Optimal Cluster Head Positioning Algorithm for Wireless Sensor Networks", Sensors 2020, 20, 3719.

[13] A. Ghosal, S. Halder, "Lifetime Optimizing Clustering Structure Using Archimedes' Spiral-Based Deployment in WSNs", IEEE Syst. J. 2017, 11, 1039–1048.

[14] G. Smaragdakis, I. Matta, A. Bestavros, "SEP: A Stable Election Protocol for Clustered Heterogeneous Wireless Sensor Networks", Techical Report; Boston University Computer Science Department, Boston, MA, USA 2004.

[15] C. Chang, J. Sheu, Y. Chen, and S. Chang, "An Obstacle-free and Power-efficient Deployment Algorithm for Wireless Sensor Networks", IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 39, 2015, pp. 795-806.

[16] W. Heinzelman, A. Chandrakasan, H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks", IEEE Trans. Wirel. Commun. 2002, 4, 660–670.

Mobile Sink Optimization for Enhancing Data Delivery in Wireless Sensor Networks

Khalid A. Darabkh[1], Enas N. AL-zoubi[2], Feras A. Al-naimat[2], and Ala' F. Khalifeh[3]
[1]Dept. of Computer Engineering, The University of Jordan, Amman, 11942, Jordan, Email: k.darabkeh@ju.edu.jo
[2]Dept. of Computer Engineering, Al-Ahliyya Amman University, Amman, 19328, Jordan
Emails: {ealzubi, f.alnaimat}@ ammanu.edu.jo
[3]Dept. of Elec. and Comm. Eng., German Jordanian University, Amman, 11180, Jordan, Email: ala.khalifeh@gju.edu.jo

***Abstract* - As of late, the improvement of savvy sensors in certifiable applications has offered footing to the movement of Wireless Sensor Networks (WSNs). In WSNs, preserving sensors' energy is a challenging issue and certainly requires extreme attention as they are limited-powered nodes. Utilizing mobile sink based approaches has been proven to be effective in mitigating the aforementioned limitation. As far as the sensor nodes are concerned, keeping track of the sink location is one of the most challenging issues that is facing them. To address this concern, this article proposes an efficient clustering mobile sink-based approach for efficient data gathering. Particularly, the sink node moves in an Archimedes' spiral path with fixed speed without any location update broadcasting. The simulation results show significant improvements over which is quite pertinent in the literature in the perspective of network lifetime.**

Keywords – WSNs; clustering; hotspot problem; sink isolation; mobile sink

I. INTRODUCTION

Recent advances in sensor networks have been made by the collaboration among micro-electro-mechanical systems (MEMS) technologies, digital electronics and wireless communications. This results in developing a low-power, low-cost, small sized and multifunctional sensor nodes [1-2]. WSNs proved significant improvements over traditional networks and provided tremendous benefits [3-4]. Sensors can be placed far from the phenomenon and this is called a sense perception [5]. In such cases, large sensors equipped with complex technologies used to sense the needed parameters [6]. It is worth mentioning that a large number of tiny sensors are deployed in the phenomenon area or close to it to sense the data, process and sent it to the central units [7].

It is worth detailing that the most directly connected works, to our work, are those proposed in [8-9]. In particular, the authors in [8] proposed the Virtual Grid based Dynamic Routes Adjustment scheme (VGDRA). It attempted to decrease the routes' reconstruction cost of the nodes for the reason of maintaining the optimal routes to the latest mobile sink location without imposing multiple mobile sinks or using super nodes, thereby extending the network lifetime and improving data delivery ratio using a single mobile sink and adhering to the low-energy theme of WSNs. The network area is divided into equally sized grids where each one has a single Cell-Header (CH) and member nodes. Cell-headers are selected based on a threshold distance from the center of the cell. They are responsible for collecting data from the member nodes in the cell and send these data to the mobile sink using a virtual construction network. In addition, cell-headers keep track of the sink location and maintain recent routes to it. The mobile sink moves on the periphery of the network area and collects data from the border of cell-headers. When the cell-header discovers a recent location of the sink, it announces it to the other cell-headers through following certain rules to maintain fresh routes to the sink and ensure the data delivery with minimal control overhead. Interested readers may refer to [8] for further details.

The authors in [9] proposed a Virtual Grid based Data Dissemination (VGDD) protocol as an extension of VGDRA. This protocol aimed to optimize the data delivery ratio and increase the network lifetime while adhering to the limited low energy supplies in the sensor nodes. Following the same basic guidelines as VGDRA, a VGDD network is divided into grids according to the number of nodes and a specific number of cell-headers participate in the routes reconstruction based on VGDRA rules. However, in VGDD, the sink moves in a counter clockwise circular path at the border cells of the network but inside the cells not as that in VGDRA. Each grid contains a CH, which is selected based on the threshold distance from the center of the grid and a threshold energy. This threshold distance is slightly increased and the threshold energy is decreased with the time to ensure the presence of such nodes. The remaining nodes in grids are participated as member nodes. Actually, CH is responsible of sending the collected data from its grid to the mobile sink if it is located in its range or to a neighboring CH to be then forwarded to the mobile sink. In addition, certain subset of CHs keep track of the mobile sink trajectory based on VGDRA rules through the CHs and gateway nodes which construct the virtual infrastructure. This approach reduced the control overhead and latency as well as improved the data delivery ratio.

In this work, a Genetic Clustering with Mobile Sink Protocol (GCMP) is proposed. in this proposed protocol and during the movement of mobile sink, the time synchronization is maintained among the mobile sink and static sensors, thereby avoiding any control overhead required. Moreover, the cluster heads are selected based on a powerful genetic algorithm in which the most influential parameters have been considered. However, the rest of paper is organized as follows. Section II details the proposed protocol. In section III, simulation results and discussions are illustrated. Finally, the work is concluded in Section IV.

II. PROPOSED PROTOCOL

A mobile sink follows an Archimedes' spiral movement through the network area of interest. It starts from the center of the network area to its borders, which is called the forward movement, and then ends in the reverse movement. While touring, the sink collects data and sends it to the administrator. The main objective of this work is to allow the nodes to determine the location of the sink without any control overhead messages. In addition, an efficient cluster

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

head selection algorithm is proposed which mainly relies on the genetic philosophy.

A) Network Model

This model is considered over a squared network area, where the area is divided into equal squared grids (cells) as shown in Figure 1. In this hierarchical-based model, the sensor nodes sense, process, transmit and receive both; data and control messages. Some nodes act as CHs while the other nodes, which are called member nodes, sense the parameters and transmit the readings to their CHs. The hierarchical-based model is a three-layer architecture, where the mobile sink presents the highest layer, CHs act as transmitters between member nodes and the mobile sink while member nodes represents the lowest layer.

Similar to [1, 3, 6], the network lifetime is divided into rounds, where the mobile sink is in charge of defining the round time. The first round contains two phases, namely, the set-up and steady phases. The next rounds may have the steady phase only or both phases. The setup phase occurs not in every round, but in every patch of rounds. In the set-up phase, the mobile sink divides the area into virtual grids, determines which nodes belong to each grid, selects CHs for each grid based on the proposed genetic algorithm, and consequently maintains neighborhood and CHs tables.

B) Division of Network Area

The structure of the virtual grid is performed based on the nodes number in the network. The reason behind this partitioning is to spread the energy consumption uniformly among the nodes. To obtain the best number of grids and cluster heads, the adopted heuristic that has been used is the same as that employed in VGDRA [8] and VGDD [9]. The network area is divided into 4, 9 and 16 grids, when the number of nodes is 100, 200, and 300 respectively based on the following formula:

$$K = \begin{cases} 4, & N \le 100 \\ 9, & 100 < N \le 200 \\ 16, & 200 < N \le 300 \end{cases} \quad (1)$$

Having that K is a squared number that presents the grids number and N presents the nodes' number. According to this, nodes can be localized in term of their coordinates to columns, rows, and grid IDs. The network area is defined as A. consequently, the grid length (cell length) and the network side length can be calculated as:

$$Cell\ length = \sqrt{\frac{A}{K}}. \quad (2)$$

$$Network\ side\ length = \sqrt{A}. \quad (3)$$

C) Sink Movement

In order to mitigate the control overhead and increase the network lifetime, the mobile sink moves according to Archimedean spiral motion to cover the network area and be close to the CHs as shown in Figure 2. At the beginning, the mobile sink is located at the center. It then broadcasts only one beacon message contains the sink location, sink spiral equation, as well as sink speed. While receiving this message, all nodes will reply in return with a beacon message, which contains their IDs and locations. As soon as the mobile sink receives the messages from all nodes, it will trigger the grids' formation, utilizing aforementioned divisions, select the CHs for each grid, utilizing the genetic algorithm discussed in the subsequent section, and ultimately send these data to all nodes. According to that, each node can determine the sink location at any time without the need of the frequent broadcasting of the current location. Below are the spiral equations.

$$Sink_y = initial\_location_y + r_i(t)\sin(\theta(t)), \quad (4)$$

$$Sink_x = initial\_location_x + r_i(t)\cos(\theta(t)), \quad (5)$$

$$r_i(t) = ut + r_0, \quad (6)$$

and,

$$\theta(t) = wt + \theta_0, \quad (7)$$

Providing that r refers to the distance from the origin, a is the initial point, b is a constant which determines the distance between arms, θ is the angle from x-axis, $r_i(t)$ is the radius of the sink location in time (i), $\theta(t)$ is the angle of the sink location in time (i), r_0 is the initial radius of the sink location, θ_o is the initial angle of the sink location, u is the velocity that determines the distance between the arms of the spiral (affects the distance between the arms), w is the angular velocity that determines the speed of the circular path of the spiral.

D) Cluster Head Selection

Due to its high reputation in finding the optimum solutions, the genetic algorithm is employed in this work to determine the appropriate nodes to become CHs for a specific period. CH replacement occurs when the energy of the current CH becomes equal or below a certain energy threshold. Therefore, a fitness function is operated which is basically parametrized by the following factors:

$$Fitness\ function = \frac{(wt_2 \times f_2 + wt_3 \times f_3)}{(wt_1 \times f_1)}, \quad (8)$$

where,

$$f_1 = \sqrt{(y_{sink} - y_{node})^2 + (x_{sink} - x_{node})^2}. \quad (9)$$

and,

$$f_3 = N_{SN} / A_{UN}. \quad (10)$$

Taking into account that f_1 refers to the distance between the candidate node and current location of the mobile sink, f_2 represents the residual energy of the candidate node, and f_3 denotes for the candidate node's density that reflects the number of neighboring sensor nodes (N_{SN}) per unit of area (A_{UN}). Lastly, wt refers to the weight of factor. The steps of selecting a CH are demonstrated below in sequence:

- The mobile sink broadcasts a HELLO message for all nodes which contains mainly the sink initial coordinates and sink speed (scheduling).
- Nodes reply with a HELLO message containing the node's ID and node's coordinates.
- The sink performs the setup phase by dividing the field into equal grids, deploying the sensor nodes utilizing uniform distribution, and accordingly selecting the CHs based on the genetic algorithm.
- As far as CHs selections are concerned, the initial population is found from the nodes taking into account that the candidates must satisfy the previous conditions which are briefed as follows: the distance between the

candidates and sink should be the minimum, the residual energy of the nodes must be the maximum and the density of the nodes should also be the maximum.

- After selecting the initial population, the fitness of each node must be evaluated using Equation 8.
- The crossover is applied to generate new seed.
- Mutation process is applied on the new seed.
- The fitness weights of the new seed will be compared with the parents to select the optimum CH.
- The genetic algorithm will be terminated as soon as the required goals are met, consequently the optimum CHs are found.

E) Radio Energy and Channel Propagation Models

Sensor nodes communicate with each other and transmit the sensed data toward the mobile sink node. The main factor of energy consumption in WSNs is communication, which depends on the distance between the transmitter and receiver. In this work, the energy model that is used is the same model as that used in [1, 3, 6]. In this proposed model, the transmitter dissipates energy to run the power amplifier and electronic transmit circuit while the receiver dissipates energy to receive the radio signals. The electromagnetic wave propagation is modeled as the power law function of the distance between the sender and receiver. Figure 3 shows the radio energy model used along with the channel propagation model. Particularly and referring to this figure and to be able of transmitting a packet of length T-*bit*, the transmitter circuit dissipates $E_{TX}(T,R)$ which basically refers to the energy depleted for transmit electronics (i.e, $E_{elec}\ x\ T$) and transmit amplifier ($\varepsilon_{amp} x\ T\ x\ R^n$) taking into account that R represents the distance between transmitter and receiver antennas. At the receiving end, there is some energy dissipated for receiving electronics ($E_{RX}(T)$) which is formulated by ($E_{elec}\ x\ T$).

III. Simulation Results and Discussions

Numerous simulations were conducted to measure the network lifetime which can be further represented by two metrics whereas all are measured in rounds. The first metric refers to the First Node to Die (FND) which exactly denotes for the time from the beginning of the network operation until the moment the first node dies. The other one represents the Last Node to Die (LND) which precisely defines the time from the beginning of the network until the moment the last node dies. The simulation setup is presented in Table I. Interestingly, Figures 4 and 5 represent the number of alive nodes versus round number considering the network sizes of 50m x 50m and 200m x 200m, respectively. In both networks sizes, the network density is kept fixed as 0.01nodes per m^2. As can be noticed from Figures 4 and 5, the number of alive nodes decreases as the round number increases in the both protocols which is to be expected as increasing the round number indicating depleting more energy. According to these figures, the performance of GCMP surpasses that in VGDD in the perspectives of FND and LND metrics. In VGDD protocol, the mobile sink moves in a circular path around the network area only. Thus, CHs consume high energy to transmit their data to the mobile sink bearing in mind that the faraway CHs may transmit their data to their neighbors' CHs toward the mobile sink or directly to the mobile sink only if the sink is located within the boundaries of their clusters. Furthermore, nodes also consume energy in electing CHs whereas sensor nodes exchange their energy levels and locations for selecting new CHs. In addition, CHs consume energy more than other ordinary sensor nodes as they are required to keep track of the location of the mobile sink and consequently reconstruct the updated routes based on sink movements. Based on the above, the energy level of sensor nodes declines every round much more than that in our proposed protocol until the network terminates.

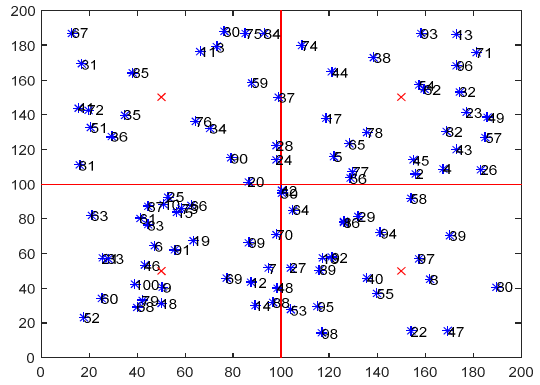


Figure 1. Network division when the number of nodes equals 100

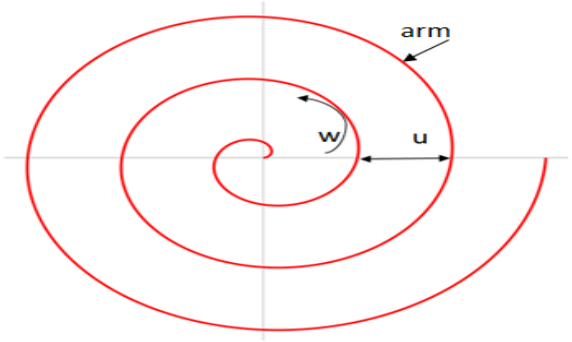


Figure 2. Archimedean spiral motion

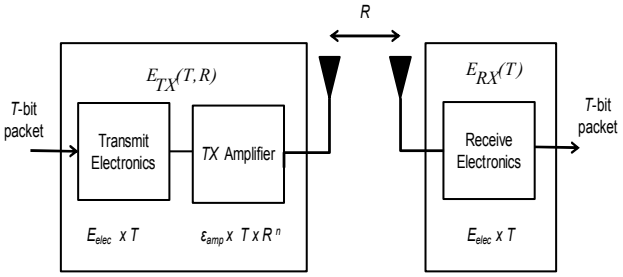


Figure 3. Radio energy and channel propagation models [3, 6]

Table I. Simulation parameters

Parameter	Value
Network size	50m x50m
Initial energy of node	0.5 J
Transmission Energy (E_{TX})	50 nJ/bit
Receiving Energy (E_{RX})	50 nJ/bit
Data Aggregation Energy	5 nJ/bit
Transmit Amplifier Energy	Free space:10 PJ/bit/m^2 Two ray: 0.0013 PJ/bit/m^4
Data packet length	6400 bits
Control packet length	200 bits
Crossover Distance	87.0 m
α	0.1
wt_1, wt_2, wt_3	wt_1: 0.3, wt_2: 0.4, wt_3: 0.3

It is noteworthy to mention that achieving a more prolonged network lifetime as in GCMP do not come out of the blue, but rather due to magnificent enhancements. Firstly, the movement path of the mobile sink is the Archimedean spiral motion. This makes the mobile sink come closer to the sensor nodes and further cover all the network area. As a result, the long-distance communications between the mobile sink and the CHs are substantially decreased, thereby consuming less energy and achieving longer network lifetime. Secondly, all CHs perfectly know the sink location at any time without reconstructing the root path and loosing energy as a result of mobile sink locations' broadcast messages required. Thirdly, the CHs electing process, which is based on genetic algorithm, offsets the need of exchanging the messages among the nodes in the clusters taking into account that the CHs selected not only should be the nearest nodes to the mobile sink, but also have the highest energy and highest neighboring sensor nodes. Fourthly, the setup overhead is not required in every round time, but instead, in the first round only. In other words, the network re-clustering occurs as soon as the energy of one of the current CHs dips below the energy threshold. If that is the case, then the CH role will rotate among the sensor nodes, based on the genetic algorithm, to evenly distribute the energy consumption.

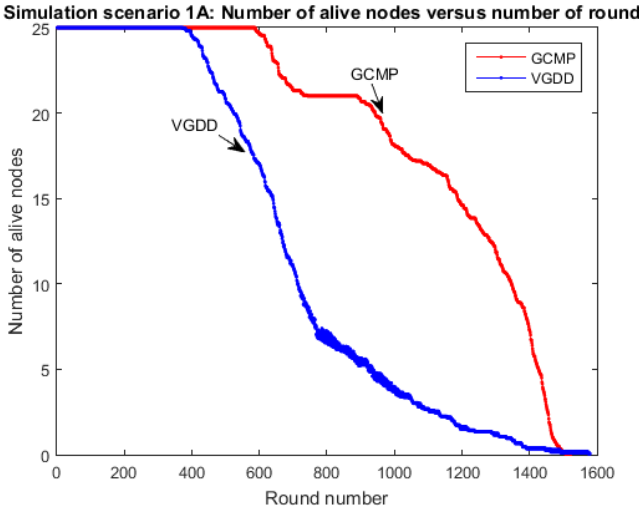


Figure 4. Number of alive nodes versus round number considering the network size of 50m x 50m

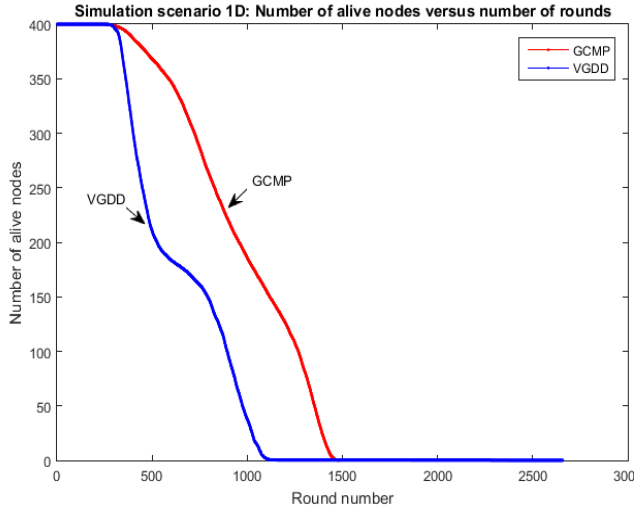


Figure 5. Number of alive nodes versus round number considering the network size of 200m x 200m

IV. Conclusion

This work presents a mobile sink optimization protocol for enhancing data delivery in wireless sensor networks, namely, GCMP. In this protocol, the time synchronization among static sensor nodes and mobile sink is maintained. As a result, the sensor nodes can guarantee finding the current sink location, whenever needed, without the need of broadcasting its current location periodically. A balance in energy consumption is achieved through the incorporation of a clustering model in which the cluster heads are selected based on an efficient genetic algorithm whereas the broadcast declarations are kept to minimum. In our genetic algorithm, the following dimensions have been taken into consideration: 1) the distance between the node and current location of the mobile sink, 2) node's residual energy of the node, 3) node's density, which refers to how many nodes surrounding the current node per unit of area. In GCMP, the mobile sink can be close to all sensor nodes and further cover the whole network area. Furthermore, it allows the interception between the mobile sink and sensor nodes in both directions (forward and reverse). Additionally, re-clustering operations do not occur every round but rather every batch of rounds, thereby maintaining low control overhead. In short, numerous number of simulations were conducted for evaluating the performance of GCMP and comparing it further with that of VGDD in the perspective of energy consumption. As a result, the simulation results of this work show significant improvements over that of VGDD protocol.

REFERENCES

[1] Khalid A. Darabkh, Noor J. Al-Maaitah, Iyad F. Jafar, and Ala' F. Khalifeh, "EA-CRP: A Novel Energy-aware Clustering and Routing Protocol in Wireless Sensor Networks," *Computers and Electrical Engineering*, vol. 72, pp. 702-718, November 2018.

[2] Manshahia, M. S., "Wireless sensor networks: A survey," *International Journal of Scientific & Engineering Research*, vol. 7, pp. 710-716, 2016.

[3] Khalid A. Darabkh and Laith Al-Jdayeh, "AEA-FCP: An Adaptive Energy-aware Fixed Clustering Protocol for Data Dissemination in Wireless Sensor Networks," *Personal and Ubiquitous Computing*, vol. 23, pp. 819–837, 2019.

[4] N. Sabor, S. Sasaki, M. Abo-Zahhad, and S.M.Ahmed, "A Comprehensive Survey on Hierarchical-Based Routing Protocols for Mobile Wireless Sensor Networks: Review, Taxonomy, and Future Directions," *Wireless Communications and Mobile Computing*, 2017, 2818542, 2017.

[5] T. Rault, A. Bouabdallah, and Y.Challal, "Energy efficiency in wireless sensor networks: A top-down survey," *Computer Networks*, 67, pp.104–122, 2014.

[6] Khalid A. Darabkh, Mohammad Z. El-Yabroudi, and Ali H. El-Mousa, "BPA-CRP: A Balanced Power-Aware Clustering and Routing Protocol for Wireless Sensor Networks," *Ad hoc Networks*, vol. 82, pp. 155-171, January 2019.

[7] Wang, J., Cao, J., Ji, S., Park, J. "H. Energy-efficient cluster-based dynamic routes adjustment approach for wireless sensor networks with mobile sinks," *The Journal of Supercomputing*, vol. 73, 3277-3290, 2017.

[8] A. W. Khan, A. H. Abdullah, M. A. Razzaque and J. I. Bangash, "VGDRA: A Virtual Grid-Based Dynamic Routes Adjustment Scheme for Mobile Sink-Based Wireless Sensor Networks," *IEEE Sensors*, vol. 15, no. 1, pp. 526-534, 2015.

[9] Abdul Waheed Khan, Abdul Hanan Abdullah, Mohammad Abdur Razzaque, Javed Iqbal Bangash, Ayman Altameem, "VGDD: a virtual grid based data dissemination scheme for wireless sensor networks with mobile sink," *International Journal of Dist Sen Nets*, vol. 11, no. 2, p. 890348, 2015.

A New Approach Model of e-Visual Career Application in Distance Education

Mustafa TUNAY
Computer Technologies Robotics Application and Research Center, Department of Computer Engineering,
Istanbul Gelisim University,
Istanbul, TURKEY
mtunay@gelisim.edu.tr

***Abstract*—A new approach for ranking visual career counselling based on distance measures is explained. In this paper, the traditional method of career counselling will be carried out by means of distance consulting services to make career counselling. Especially, professional education in the capacity limitation and professional working demand for education has led to a proliferation of virtual applications. Hence, e-Visual Career Application has been developed as a web-based program for clients who have taken professional counseling efficiently. With this application, it will help clients in the process of discovering and making decisions to their interests, abilities, personalities, skills and values. This paper relies on the implications of introducing online visual career consulting in the login sessions and how it affects the people (supervisor and client), the processes (consulting), and the organizations involved.**

Keywords—Visual Career Counselling, Traditional Method, Virtual Application, e - Visual Career Application

I. INTRODUCTION

This career application system is done with specially developed software for providing more fruitful advising system on career development by considering interests, capacities, personalities, abilities and values that are developed in order to help people while they are discovering the right track and making decisions on their career development in this paper.

Distance education [1-2] is not a new concept and "Distance education", which eliminates the obligation of the instructor and the student to be in the same place; it is an education system, it is an alternative system to today's formal education system and its success rate increases day by day with the developing technology. Although a definite starting date cannot be determined but the use of distance education applications in an organizational structure is encountered in the 19th century.

In recent years, information and educational technology in distance education [3-5] have developed, while more information has been produced than in the past. These developments in distance educatrion changed the lifestyles of people, their needs and thoughts in obtaining information. As a result of the changes that occur in parallel with the development of educational technology in distance education, the most educational institutions is affected from globalization. Technological tools [6-7], which are always with individuals, increase the importance of the use of e-learning tools [7-10] for institutions and individuals to have an effective and productive structure in their career development. Thus, developing educational technology, reaching the information from the reliable sources and stating the references of the sources correctly can be considered as an important element in the development of societies [11]. For this reason, internet and computer are the most of the important communication tools which help people in reaching the scientific information on the right time.

The online education [12-14] system is a very important education system in order to prevent formal education that cannot be processed fluently, especially as a result of the measures taken against global problems. Thanks to online education, students can continue their education without delaying their courses for lost opportunity in face-to-face learning [15-17] and time and can repeat the topics covered on the system. In this regard, online education are conducted in 2 ways; asynchronous and synchronous [18-19]. Although distance education is called asynchronous education, but universities have unique learning management system (LMS) that can reflect their own characteristics and meet the requirements of the distance education applications [20] in today's world, adapt rapidly to technological innovations will arise in the medium and will provides synchronous education opportunities integrated into this system for a long term. Thus, this system as educational programmings [21-22] provide an interactive communication between instructors and students. Thanks to such educational programmings, it will be a important tools [23-24] that help students especially their career development skills.

Discovering the existence of new science, gaining new facts, and using that information in the right place and on the right time lead to new job areas (career development skills) for people in the world. It can be seen that most of the people from different profession needs to update themselves and their knowledge each passing day. Providing counselling services to the people who are in the working life actively will help them to address the need of the people around us and this will raise their activities in their profession. However in recent days, it can be seen that lack of counselling advisors cause the problems in the processing of the duties. For this reason, applying the use of “distance web-based advising career application” has a huge importance. This application is dealing with management of synchronous, asynchrony [19] and the mix of them when supervisor and client conducting the consulting service at the same time (synchronous), when the advisor and client learn and use the counselling material separately (asynchrony) and hence the application system as learning management system [25-29] that use both of them together is called as mixing. e-Visual Career Application

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

(e-VCA) System is a kind of distance web based for handling most of the career consulting activities as it can be done with face-to-face education. It is possible to have audio-visual communication with thorough independent career development advising career application system is also look like a face- to-face environment. Client may follow the notes of the counselling weekly, personal information forms, tests, measurements, inventory, technics and weekly research activities through this career application system, they may interact with their supervisors and they may communicate in login sessions. The scope of this career application system is to gather requirements, analyze, design, and develop the e-Visual Career Application website. The objective of this engagement creates e-Visual Career Application (e- VCA) for supervisors and clients that will achieve the following goal:

- Supervisors will be able to do registration of the career application system by using on the internet.
- Clients will be able to attend lively meeting arranged before.
- A client registered in the system application will be able to access the counseling materials easily from their portal through this career application system.
- All the questions are raised by the client will be able to produce chain of discussion have a change to discuss with his/her supervisor through the career application system.
- Clients will get chance to ask for an extra meeting from their supervisors.
- This career application system consists 5 sessions and each session is 45-50 minutes in a week.
- Client will be able to do the personal information form, tests, measurements, inventories, techniques and weekly searching tasks on the career application system.
- The attendance of clients will be able to be followed by the advisors with the help of career application system. The supervisor will be able to arrange an extra meeting hourly and will be able to send the counseling materials on the web.
- This application system will be the center of most of the developmental areas related to teaching and learning with the raising application areas, with the harmony of the learning strategy, without any problem related to place and timeframe and with the flexible learning methods.
- With the scope of the career application system, the content of the counseling is prepared for five week and all the counseling materials will be opened for supervisor during the year.

1.1 Executive Summary

Implementing the system application is using common internet technologies with PHP and Mysql as the backbone, provides staff with a tool to easily communicate, distribute information and facilitates collaboration across the entire organization.

1.1.1 Technologies Used

This career application system is a kind of web application and is developed as;

1. Database Design (MySql)
2. Coding (PHP)

1.1.2 Browser Compatibility Statement

We request you to use the career application system in the latest version of browsers and are available from the market. Because the older versions may not support and some design part of the career application system needs manual update or automatic update to keep up to date of web browser.

1.1.3 Software Interface

The purpose of this section is to define all online screens that are part of the career application system and their interdependencies and to provide details on the functionality of the screen based on a user's actions.

1. Web Server: .NET Framework, Web Browser.
2. Database Server : MySQL

1.1.4 Security of System

PHP (coding) is a fast and easy-to-learn language, for this reason, we can cause software-related security vulnerabilities. The proposed system data is protected with password-based logins by SSH key, or secure shell that is an encrypted protocol used to administer and communicate with servers. SSH Key is any kind of authentication, including password authentication and it is completely encrypted. Thus, SSH keys use encryption to provide a secure way of logging into your server and are recommended for all users. For SSH key authentication is showed in Figure 1. There is a important point for user must place your public SSH key on the server in its proper directory.

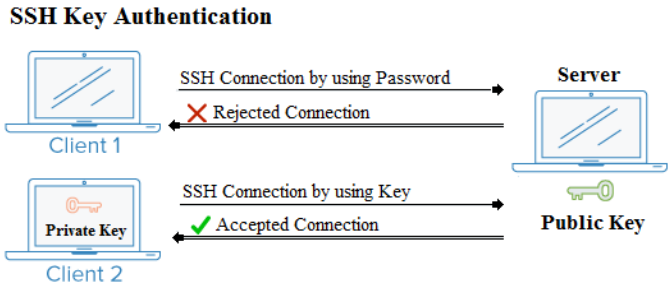


Fig. 1: Security of System by using SSH Key

According to SSH keys Authentication; a private and public key pair are created for the purpose of authentication. When password-based logins are allowed, malicious users can repeatedly attempt to access a server, especially if it has a public-facing IP address.

II. THE APPLICATION SYSTEM

In a distance education process, there are many Learning Management System (LMS) applications provide learning content presentation in a web based learning as web server software (MySQL, Apache, Oracle e.g.) for performing learning content that can be used in many different phases as career application systems. The significant time required to develop scales has also been a consistent finding in studies involving distance education in career and technical education career application system. For this reason, e -Visual Career Application (e-VCA) *System* is a kind of distance web based for handling most of the career consulting activities. In Figure 2 shows that the details regarding the e-Visual Career Application (e-VCA), scales (files) and web server software are used in a career application system.

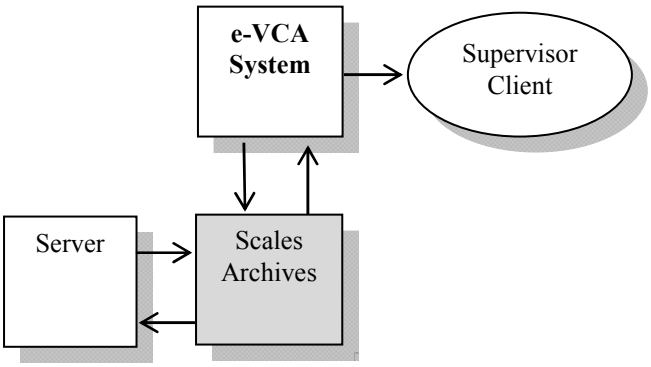


Fig. 2: e-Visual Career Application (e-VCA) and Web Server Software

The career application system is designed and represented the general scope with a contextual boundary of e-VCA System. It describes the main objective of the system and its entities involved functionality. This career application system is designed totally five steps to show how it is easy to start using its organization functionality such as; Login, the page of Home, Archives and Scales, Clients' Information and Sessions.

2.1 Login

The first stage function is designed the page of login that is expected to play an important role in the conducting of online career counseling. The purpose of the login page is first required to register by from clients on the online website. This page also has access of the system that is secured by the login/password mechanism (by using private e-mail address) for the login id is created an account from the clients at the time of registration and provided it from their supervisors. Figure 3 shows that the details regarding the page of login in a career application system.

2.2 The Page of Home

The second stage function is designed about view of registration information. The purpose of the home page has demonstrated views of total approved and unapproved clients are available. It has included how many total/active topics as personal information form, tests, scales, inventories, and research techniques weekly that are presented. In Figure 4 shows that the details regarding the page of home in career application system.

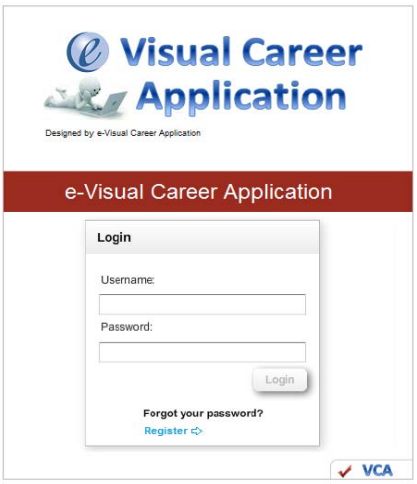


Fig. 3: The page of Login

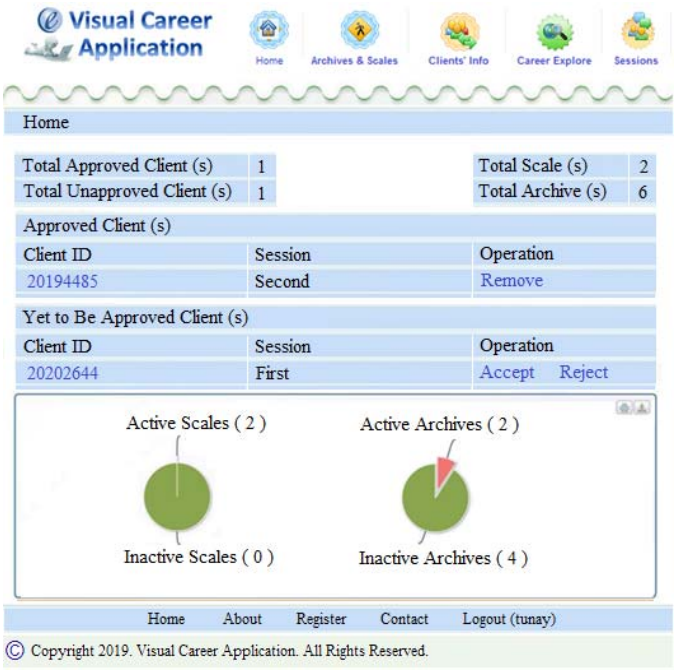


Fig. 4: The page of Home

Next one is about the lists of approved clients; it has been viewed with their information and supervisors can manage the client's information such as their control by hitting of remove link, the listing their information that are showed. Secondly, this part has two important roles as print option; it is about print out of summary for active and inactive documents. And next one is about chart; it also has option to download and view in various formats about clients' information. The detail regarding the page of Home is given in Figure 4.

2.3 Archives and Scales

The third one is designed the page of Achives and Scales; it has all privileges to edit/delete all the files /directories created by supervisor who uploads topics (personal information form, tests, scales, inventories, and research techniques) and clients can view all the files uploaded by supervisor according to their sessions. The detail regarding the page of Archives and Scales is given in Figure 5. Supervisor can manage all of information as sending new tests, scales, inventories and new sources for special member clients. They also can search any member of client and take information about searched the member client.

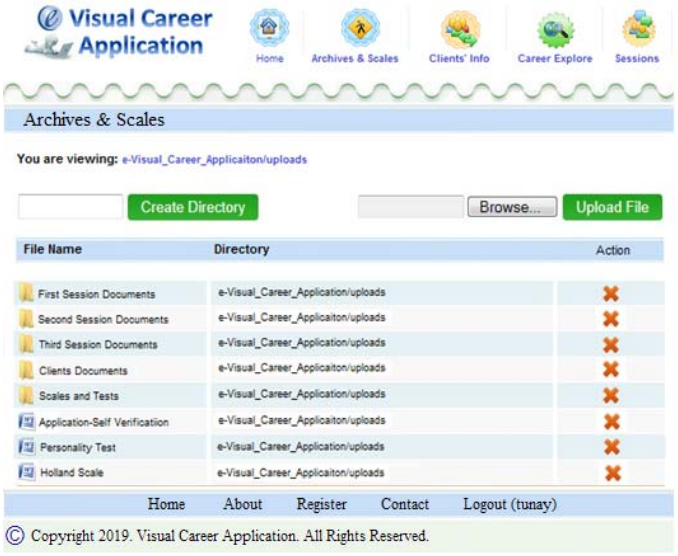


Fig. 5: The page of Archives and Sclaes

2.4 Client's Information

The fourth stage is designed for the page of Client info; it allows view of clients' situation (personality information, address, attending session and other information). It has two important roles as adding clients and editing clients.

2.4.1 Adding Clients

To add a client, go to the clients tab and fill in the form and click the add client button at the bottom of the page. Newly-added clients are automatically given active status, allowing them to sign-in to e-Visual Career Application immediately. However, newly-added clients must be told how to access e-Visual Career Application System that sent info feature to approved e- mail by using the activation with client information as username, password, and URL to sign-in.

2.4.2 Editing Clients

To edit a single client, click on his/her row, update their information, and click on the save changes button at the bottom of the page. To edit alter multiple clients at once by clicking on the checkbox next to their name (or select all by clicking on the checkbox in the header) in combination with the button at the bottom of the page.

Moreover, it gives change to supervisor can manage to keep certain clients from signing- in to e-Visual Career Application sets their account status as active or inactive. The detail regarding the page of Search is given in Figure 6.

Fig. 6: The page of Search

2.5 Session

The last one is designed for the page of Session. Career Counselling consult with scales determined by the consultant to implement the system for the distance career education of the necessary personal information, forums, testing, inventory, and make appointment online interviews with five sessions and each session is determined as 45-50 minutes. In this context, e-Visual Career Application system offers online consultations that are relevant to the topic, inventory, scales (Holland or MBS), activities and the session on the techniques during last for five weeks. These sessions are as follows;

1. Session: Configuration and discovery.

2. Session: Scale and tests, and provide information about the application.

3. Session: Transmission of test results, and review of the advisory tasks.

4. Session: Duties discussion, brainstorming, information, research, and deal with the action plan.

5. Session: progress review, and if necessary to determine new goals.

At the same time, the client requests extra appointments to take any time from their supervisor by means of e-Visual Career Application system, session on the topic which is applicable to the person receiving vocational career education, carried out sharing inventories with supervisor electronically.

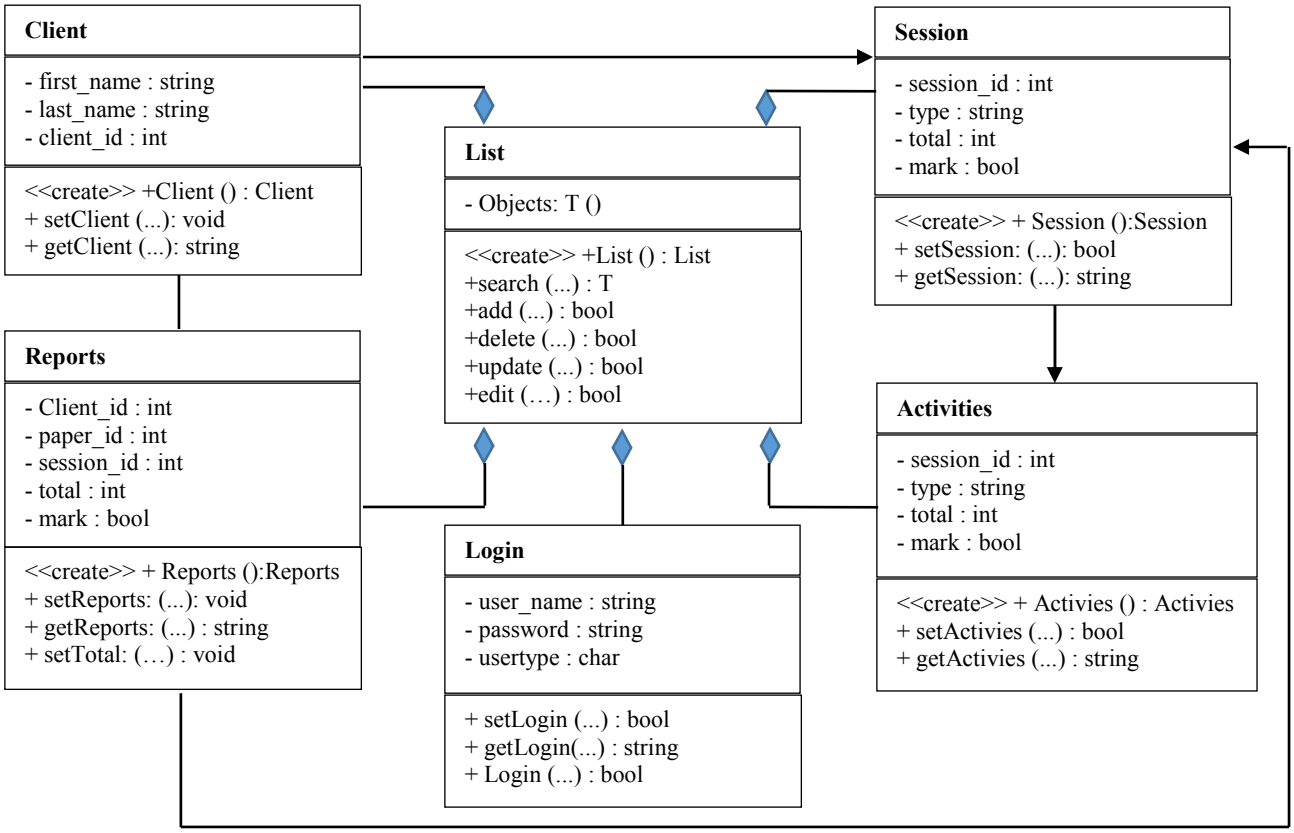


Fig. 10: The diagram model of the proposed system

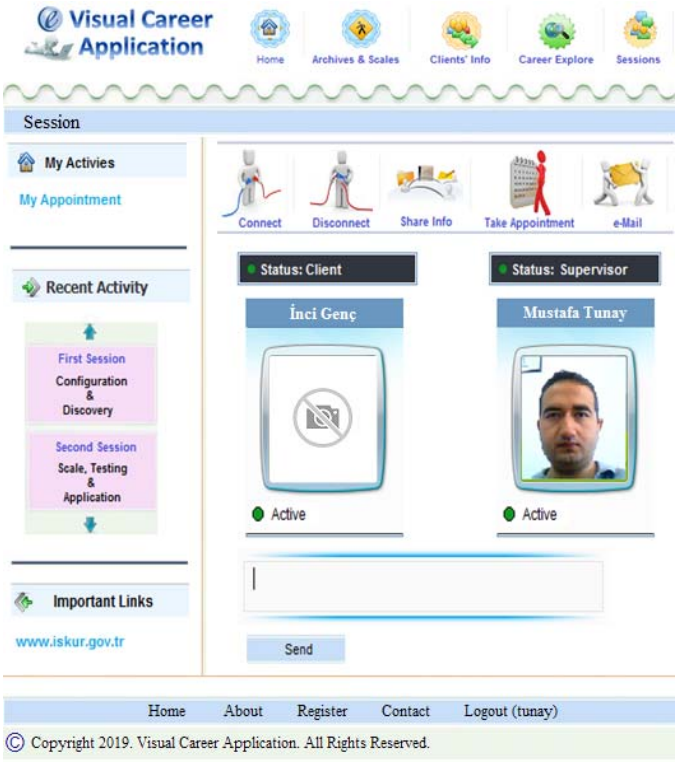


Fig. 7: The page of Session

This application system is applied on a blind student's İnci GENÇ at Eastern Mediterranean University in TRNC. She got the necessary personal information, forums, testing; inventory on the application system by using record voice, and make appointment online interviews with five sessions and each session was as 45-50 minutes with his supervisor. She got a great achievement at the end of five sessions. The detail regarding the page of Session is given in Figure 7.

This application system is ready to apply on close to graduate students (this semester) at Istanbul Gelisim University in Turkey for improving their career skills.

III. The Model Application System

The model of career application system is designed with supervisor, client and functional activities are described. In Figure 8 represents what are the bounders and scope of e-Visual Career Application system. It describes the main objective of the system and is entities involved. Moreover, the diagram represents the general scope with a contextual boundary of e-Visual Career Application (e-VCA) system describes the main objective of the system and its entities involved. All details regarding the context diagram of proposed system is given in Figure 9.

This career application system is done with specially developed software for providing more fruitful advising system on career development by considering interests, capacities, personalities, abilities and values that are developed in order to

help people for improving career skills in their future life and is figured out the overall system with its entities.

The model of e-Visual Career Application used sequence diagram generally and figure out the overall system. This model of diagram are interactions between a system's clients-system and supervisor-system with among objects in the system when relations occur with whole application system in Figure 10.

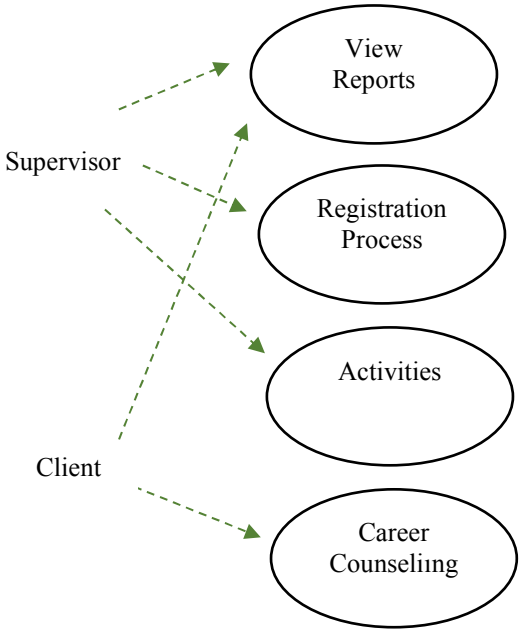


Fig. 8: The Model of e-Visual Career Application (e-VCA) System

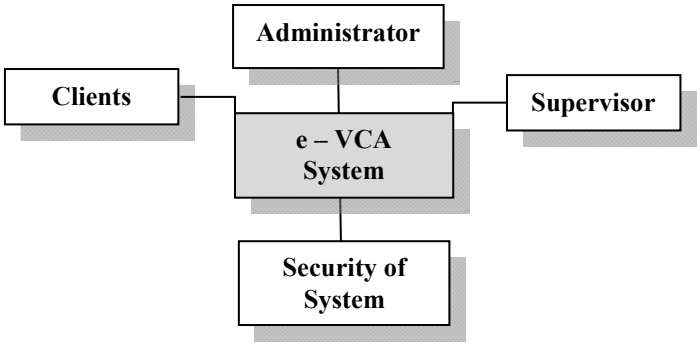


Fig. 9: The context diagram of e-Visual Career Application System

IV. CONCLUSION

This paper proposes the e-visual career application system is designed for clients in helping to take professional career counseling efficiently and will help clients in the process of discovering and making decisions to their interests, abilities, personalities, skills and values. The designed career application system is included many features that improves pedagogical scales, inventory and many needed tools for of the academic disciplines, developing career in distance education programs utilize technology to a high degree. In addition, we can say that it is a system which is very useful and using time period for the specialists in terms of their distance, for the use of disabled people and new graduated students to improve their career skills in future life.

In the future, this study will need to be conducted to determine the most effective ways to utilize distance education in the delivery of professional career counselling.

ACKNOWLEDGMENT

The author would like to thank *Computer Technologies Robotics Application and Research Center,* Istanbul Gelisim University, Turkey.

REFERENCES

[1] Distance Learning Resource Network. What is distance education? Retrieved April 23, 2003 from www.dlrn.org/library/dl/whatis.html.

[2] Oliveira, W. and Do Amaral, L. A. M. (2019). "Distance Education as a Service System," 2019 IEEE World Conference on Engineering Education (EDUNINE), Lima, Peru, 2019, pp. 1-6, doi: 10.1109/EDUNINE.2019.8875822.

[3] Flowers, J., (2001). Online learning needs in technology education. Journal of Technology Education, 13 (1), 17-30.

[4] Hillesheim, G. (1998). Distance learning: Barriers and strategies for students and faculty. Internet and Higher Education 1 (1), 31-44.

[5] Kanakana-Katumba, M. G. and Maladzhi, R. (2019). "Online Learning Approaches for Science, Engineering and Technology in Distance Education," 2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Macao, Macao, 2019, pp. 930-934, doi: 10.1109/IEEM44572.2019.8978892.

[6] Tesone,D. V., Kelsey,R. L., Gibson, J. W. and Blackwell, C. W. (2002). "Using multimedia tools to enhance communication in the online classroom," *Proceedings. IEEE International Professional Communication Conference*, Portland, OR, USA, 2002, pp. 199-206, doi: 10.1109/IPCC.2002.1049103.

[7] Eroglu, M. A. (2009). "Academic document sharing at IAU members and distance education tools (samples of Turkish language)," *2009 International Conference on Application of Information and Communication Technologies*, Baku, 2009, pp. 1-5, doi: 10.1109/ICAICT.2009.5372589.

[8] Paulsen, M. F.; Keegan, D.; Dias, A.; Dias, P.; Pimenta, P.; Fritsch, H.; Föllmer, H.; Micincova, M. & Olsen, G., 2002. Web-Education Systems in Europe, Hagen: Zentrales Institut für Fernstudienforschung, Fern Universität.

[9] Vargas, R. P., Hortelano,M. R., Abad, L. T., Carrillo J. C., and Berlinches, R. H. (2018). "Teaching cloud computing using Web of Things devices," *2018 IEEE Global Engineering Education Conference (*EDUCON), Tenerife, 2018, pp. 1738-1745, doi: 10.1109/EDUCON.2018.8363444.

[10] Angles,R., Silvestre, L. and Pincheira, G. (2019). "Edutrivias: Web support system based on trivia for the teaching-learning process," *2019 38th International Conference of the Chilean Computer Science Society (SCCC)*, Concepcion, Chile, 2019, pp. 1-8, doi: 10.1109/SCCC49216.2019.8966392.

[11] Paloff, R., & Pratt, K. (1999). Building Learning Communities in Cyberspace. San Francisco: Jossey-Bass.

[12] Kanakana-Katumba, M. G. and Maladzhi, R. (2019). "Online Learning Approaches for Science, Engineering and Technology in Distance Education," *2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Macao, 2019, pp. 930-934, doi: 10.1109/IEEM44572.2019.8978892.

[13] Shadmanova, G., Mirzaev S. and Karimova, X. (2019). "Upgrading Higher Education through Open Online Courses," *2019 International Conference on Information Science and Communications Technologies (ICISCT)*, Tashkent, Uzbekistan, 2019, pp. 1-3, doi: 10.1109/ICISCT47635.2019.9012019.

[14] Zhu, J. (2010). "A hybrid online-education strategy for delivering engineering and technology courses," *2010 International Conference on Networking and Digital Society*, Wenzhou, 2010, pp. 448-451, doi: 10.1109/ICNDS.2010.5479464.

[15] Swan, M., & Jackman, D. (2000). Comparing the success of students enrolled in distance education courses vs. face-to-face classrooms. Journal of Technology Studies, 24 (1). Retrieved March 11, 2003 from http://scholar.lib.vt.edu/ejournals/JTS/Winter-Spring-2000/swan.html.

[16] Merenkov, D. V., Shirinskii, S. V., Korkin, V. S., Zhokhova, M. P., Osipkin, S. V. and Dergachev, P. A. (2020). "E-Learning Resource for Electric Machine Design," 2020 V International Conference on Information Technologies in Engineering Education (Inforino), Moscow, Russia, 2020, pp. 1-4, doi: 10.1109/Inforino48376.2020.9111745.

[17] Talib, A. M. and Alomary, F. O. (2018). "An Interactive Distance Education and E-Learning System Based on Multi Agent System Architecture," 2018 1st International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, 2018, pp. 1-6, doi: 10.1109/CAIS.2018.8441984.

[18] Miller, W.W., & Webster, J. K. (1997). A comparison of interaction needs and performance of distance learners in synchronous and asynchronous classes. Paper presented at the American Vocational Association Convention, Las Vegas, NV . (ERIC Document Reproduction No. 415 411).

[19] Timofeev, A. V., Tokarev, N. V. and Petrova, K. A. (2017). "Implementation of additional training profiles «LETI minor»: Using online courses," 2017 IEEE VI Forum Strategic Partnership of Universities and Enterprises of Hi-Tech Branches (Science. Education. Innovations) (SPUE), St. Petersburg, 2017, pp. 147-148, doi: 10.1109/IVForum.2017.8246076.

[20] R. D. Costa, G. F. Souza, T. B. Castro, R. A. M. Valentim and A. P. Dias, (2020). "Identification of Learning Styles in Distance Education Through the Interaction of the Student with a Learning Management System," in IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, doi: 10.1109/RITA.2020.3008131.

[21] Alexandrovna, S. T. (2018). "The Role of Distance Educational Technologies in the Development of Educational Programmes in a Network Form: Experience and Relevant Problems," 2018 IV International Conference on Information Technologies in Engineering Education (Inforino), Moscow, 2018, pp. 1-4, doi: 10.1109/INFORINO.2018.8581847.

[22] Simonson, M., Smaldino, S., Albright, M., & Zvacek, S. (2003). Teaching and learning at a distance (2nd ed.). Upper Saddle River, NJ: Merrill Prentice Hall.

[23] Zirkle, C., & Shoemaker, H. (1999). Indiana State's multiple delivery approach: Integrating industrial technology education with educational technology. The Technology Source. Retrieved March 27, 2003, from http://www.mivu.org/teaching/techsource/

[24] Zirkle, C. (2000). Preparing technical instructors through multiple delivery systems: A working model. T.H.E. Journal 28, (4), 62-68.

[25] Aydın, C., C.; Tirkes, G. (2010). Open Source Learning Management Systems in E-Learning and Moodle, Atilim University, Ankara: Turkey.

[26] Perales, M., Pedraza, L. and Moreno-Ger, P. (2019). "Work-In-Progress: Improving Online Higher Education with Virtual and Remote Labs," 2019 IEEE Global Engineering Education Conference (EDUCON), Dubai, United Arab Emirates, 2019, pp. 1136-1139, doi: 10.1109/EDUCON.2019.8725272.

[27] Szőke, J. and Kecskés, P. (2017). "Virtual proximity in distance education — The usage of a three-dimensional software," 2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), Debrecen, 2017, pp. 000529-000529, doi: 10.1109/CogInfoCom.2017.8268291.

[28] M. Maravanyika, N. Dlodlo and N. Jere, "An adaptive recommender-system based framework for personalised teaching and learning on e-learning platforms," 2017 IST-Africa Week Conference (IST-Africa), Windhoek, 2017, pp. 1-9, doi: 10.23919/ISTAFRICA.2017.8102297.

[29] Yun Wu, W., Weng, K., Wen, M., You, B. and Lian, J. (2017). "The Construction of Web-Learning System and Learning Performance Based on Online Interaction and Self-Regulation Theories-Architectural Design Curriculum," 2017 *International Conference on Information, Communication and Engineering (ICICE)*, Xiamen, 2017, pp. 288-291, doi: 10.1109/ICICE.2017.8479001.

Evaluating State of Web Accessibility Standards in Top Airline Websites

Mukesh Pant
Tata Consultancy Services
Crawley, UK
mukesh.pant@tcs.com

Abhirup Sinha
Tata Consultancy Services
Atlanta, USA
sinha.abhirup@tcs.com

***Abstract*—A study is carried out to evaluate world's top 10 Airline websites using automated web accessibility assessment tools and verify the conformance of results using manual evaluation. Collected data is normalized and standardized using statistical methods. Analyzed data is plotted in X-Y scatter plots to identify patterns. Results depict that a few websites have only shown better performance compared to others. Comparison of evaluation results shows a variation of test criteria implemented by different tools. Airline websites need better implementation of accessibility standards for digital inclusion of wider customer base. Accessibility evaluation standardization in available tools will help developers and designers to fix accessibility issues more effectively.**

Keywords—Accessibility, travel industry, airlines Accessibility

I. INTRODUCTION

Web accessibility means that people with disabilities can perceive, understand, navigate, interact with and contribute to the Web regardless of age or ability [1]. The fundamental concept of Web is based on inclusive access and reach of information. Ensuring accessibility online for all sections of users including people with disabilities is therefore a great concern. Legally and ethically, all businesses must ensure that all interfaces available for public must be accessible to disabled individuals. Failing to provide accessible user interfaces raise serious legal liability issues.

Compliance is more desirable for businesses like AirTravel websites where a rapid growth of ~50% in passenger number is recorded during last decade and it is projected to be doubled again in next 20 years. According to an agency for statistical data analysis of aviation data AIOC, 20% of 4.54 billion passengers are either elderly or disabled. That suggest 1 of every 5 customer has special needs to operate website and other online resources. This assistance is provided by following accessibility guidelines for web content.

Web Accessibility evaluation is conducted by making these accessibility guidelines as the basis. It is a procedure to verify the conformance of a website with the guidelines set by the standards organizations. Several accessibility guidelines have been developed and used to implement and enforce accessibility. Most widely used among those are Web Content Accessibility Guidelines (WCAG). WCAG 2.1 are the latest web accessibility standards for website accessibility. Various research conducted in public websites indicate that many websites are still not accessible to people with disabilities.

To develop long lasting relationships with the customers, the companies should try to follow strategies which make the customers engaged with the. Considering a large number of website users need special attentions for accessibility of website content, it is imperative to be compliant. To check Web Accessibility compliance, evaluation can be performed either using automated evaluation tools or manually by different group of people having some form of disability. Their assessment is subjective because different users have different opinions about the same issue. It is also worthwhile to note that, a failure to perform a task is not always an accessibility problem.

User testing is constrained and not scalable due to the availability of experienced users with disability. Web designers, developers and quality engineers are therefore trained as specialized accessibility experts, to perform manual or automatic evaluations. For manual evaluation, accessibility expert tries to manually inspects Web pages to find out problems which would be significant issues for users with disabilities. In this way, manual testing complements actual user testing, but the process is very cumbersome and time consuming. In addition, the expert might have potential bias while evaluating the accessibility quality of a Web page. To resolve this problem, software tools are developed to automatize evaluation to verifies conformance with guidelines. Software usually minimizes the need for human intervention. Moreover, automatic testing is scalable and objective. However, automatic evaluation has its own limitations due to programmatic limitations and subjective nature of guidelines. Automatic evaluation may not cover the same depth as manual evaluation, nor can it be as complete. Due to these limitations, we have used manual as well as automated evaluations for this study.

Present research focuses on Web accessibility evaluation of major global airlines' website. It is an attempt to access the level of compliance of top 10 global airlines from web accessibility standpoint. Automated evaluation using 4 majorly used web accessibility tools is done for the study. Also, we have done manual validation on subset of success criteria as accessibility analysis could not be completed without manual interventions. Statistical analysis of observations is done by normalizing the results to eliminate structural and configurational differences. Results are plotted on appropriate scale to identify patterns.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

II. LITRETURE SURVEY

There are numbers of case study and evaluation conducted on measure of accessibility compliance on airline websites across the world. Our literature study was focused mainly on airline websites however not limited to that.

An evaluation was done on Airlines Websites of India [2], by Gaurav Agrawal , where it has been explained how Indian Airline's websites are good in Accessibility and Usability. A case study [3] by Saleh Alwahaishi; states web accessibility compliance of Arabian Gulf Airlines websites. A conference paper [4] by Ishaq states about web accessibility analytical study conducted on Nigerian Airlines Websites and this analysis was done based on WCAG 2.0. A case study [5] completed by Abid Ismail and K S Kuppusamy states that how important Accessibility is what steps need to be taken to make websites more accessible.

Namkung, Shin and Yang (2007) [13] has stated that the iinternet is a fast-growing advertising mechanism for hospitality industry. According to the study by O'Connor and Frew (2004) [14] and Toms and Taves (2004) [17] , websites can benefit guests as this medium provide a direct communication with suppliers and help users to search and purchase products and services without any geographical or time constraints. Users are enabled to customize their own combination of products and services using web interfaces [18]. On the other hand, web platform enables suppliers, such as airlines and hotels, to better understand customer behavior and needs and can enjoy the benefits of lower distribution costs, a direct communication with less competition resulting to higher revenues (O'Connor, 2004 [15]; Wilson & Abel, 2002) [16].

To access these E-services, people with disabilities use computers with specialized software commonly called assistive technologies. Screen readers are the most popular type of assistive technology for users with visual impairments [6]. Those with hearing impairments, cognitive disabilities and motor skill impairments may require other technologies, such as voice browsers, special joysticks or trackballs.

Ggenerally used guideline for evaluating web accessibility involves assessment using WCAG 2.1 guidelines. These guidelines and success criteria are organized around four principles to allow access and use of Web content [7]:

- Perceivable – Website's user interface and all information in that website must be presentable to all type of user so that they can aware of all thing listed there.
- Operable – Websites should be operable for all type of users using their own way of operation.
- Understandable – All information should be understandable to all users.
- Robust – Content of website should be accessible even technology gets enhanced.

Automated accessibility evaluation tools are used to check whether or not the Web pages follow the WCAG 2.1 guidelines. The most popular automatic tools include Achecker and WAVE toolbar. Google has developed a complete framework called lighthouse using a popular automated accessibility tool AXE and integrated in most popular web browser Chrome. Tennon.io is another accessibility analysis tool that is gaining popularity in automated evaluation of accessibility of web applications. These tools differ in their criteria from efficiency to conformance levels (A, AA and AAA). In WCAG, the success criteria for level A are easy to meet and do not affect the website design or structure. On the other hand, levels AA and AAA are more strict and require more work [8].

According to AkgUL and Vatansever [9], human judgment is needed to provide an accurate evaluation of Web accessibility. Automatic tool evaluations, for example, cannot give a full picture of the interaction between Web contents and assistive technology; they cannot detect all violations and can therefore result in false positives and false negatives [10].

III. Data Analysis

A. Data collection

We have analyzed data for 10 most used international Airlines websites. Four most used accessibility evaluation tools are used to collect data for automated evaluation.

Data collected using AChecker, WAVE, Lighthouse and tennon.io for chosen 10 websites is used to overcome any drawbacks of a single tool use. We used AChecker [11] to automatically evaluate the home page of selected Airline websites. The tool classifies the recognized problems into the following: known problems (these are certain accessibility barriers), likely problems (these are probably accessibility barriers), and potential problems (these need a human decision). In order to reduce complexity and provide accuracy, we have only considered the known problems detected for WCAG 2.0 with a level.

The second tool used in our evaluation is WAVE [12]. WAVE is utilized to find the errors responsible for failing WCAG2.0 guidelines. WAVE also provides ARIA elements, Structural elements and Content issues present on the webpage. For this study only errors are taken into count.

Lighthouse is another popular tool being used to evaluate multiple usability features including accessibility. Using WCAG2.0 standards for AA compliance, considered 10 airline website pages are evaluated with the help of Lighthouse. We have taken the level in percentage provided by Lighthouse as a major of compliance for website. As all other automation tools are used to take error count in consideration, we have transformed the compliance percentage into error percentage by reducing it from 100. This way we have considered non-compliance factor from Lighthouse data.

Tenon.io is another tool which gave number of failures for each website, that is taken into consideration. As described in many studies in the field of Accessibility evaluations, automation tools are limited in their ability of evaluation, therefore we have considered manual evaluation of carefully selected set of criteria to access compliance of these websites. Manual evaluation of all 10 websites is done and error score is quantified to consider along with the numbers obtained using automation tools. These 10 airline websites are American Airlines (AA), United Airlines (UA), Lufthansa, British Airways (BA), Emirates, Qantas, Cathay

Pacific (CP), KLM Royal Dutch (KLM), Delta and Virgin Atlantic Airlines (VAA).

B. Statistical analysis

The range and format of data collected using automation tools and manual testing was not aligned and therefore analysis based on that data was not possible. So, we have used normalization methods to normalize the data for all tools. Min-Max normalization and z-score based (using mean and standard deviation) standardization is used to align the data on same scale.

To normalize the data at a scale of 0 to 1, min value (X_{min}) and max value (X_{max}) is calculated for each series of data. Series is normalized using below formula (1) applied to each value:

$$Xnew = \frac{X - Xmin}{Xmax - Xmin} \qquad (1)$$

For getting z-score, mean and standard-deviation for each series of data is calculated using below formulas.

$$\mu = 1/N \sum xi \qquad (2)$$

For standard-deviation:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=0}^{N}(x_i - \bar{x})^2} \qquad (3)$$

z-score for standardization is obtained using formula

$$z = (x - \mu) / \sigma \qquad (4)$$

Calculation for above normalization and standardization is done using python scripts. That will help in reusing the method and programs for further analysis of larger datasets.

Table 1: min-max normalized data for all methods

Airline	Analysis data				
	Lightho use	Wave	AChecker	Tenon.io	Manual
AA	0.000	0.063	0.044	0.138	0.286
United	1.024	0.370	0.293	0.377	0.857
Lufthansa	0.714	0.375	0.350	0.304	0.571
BA	0.381	0.208	0.012	0.524	0.571
Emirates	0.381	1.000	1.000	0.486	0.571
Qantas	0.667	0.508	0.401	1.000	0.571
Cathay Pacific	0.857	0.213	0.560	0.759	1.000
KLM	0.190	0.699	0.006	0.945	0.143
Delta	0.095	0.000	0.000	0.033	0.000
VAA	0.167	0.051	0.018	0.000	0.143

Normalized datasets are presented in table above. Table (1) shows the data for min-max normalization while Table (2) depicts the normalized data using mean and standard deviation.

Table 2: z-score data for all methods

Airline	Analysis data				
	Lighthouse	Wave	AChecker	Tenon.io	Manual
AA	-0.345	0.183	0.120	-0.426	-0.224
United	2.603	1.151	0.879	0.242	1.543
Lufthansa	1.712	1.169	1.055	0.037	0.660
BA	0.752	0.639	0.023	0.653	0.660
Emirates	0.752	3.140	3.040	0.547	0.660
Qantas	1.575	1.589	1.210	1.986	0.660
Cathay Pacific	2.123	0.658	1.697	1.311	1.985
KLM	0.272	2.191	0.003	1.831	-0.666
Delta	-0.070	-0.016	-0.014	-0.720	-1.108
VAA	0.135	0.147	0.042	-0.813	-0.666

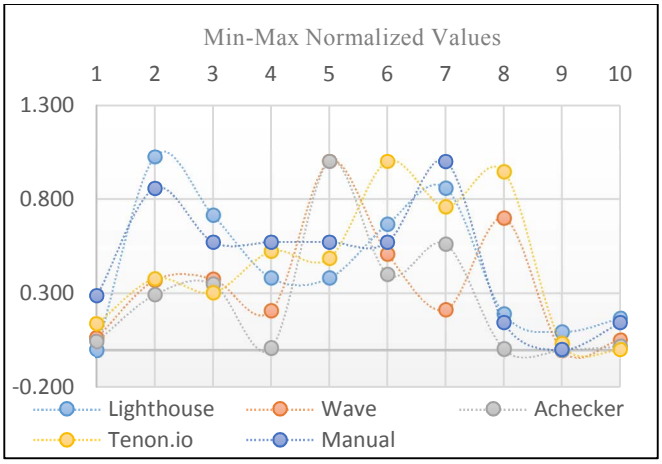


Fig. 1: Min-max Normalized plot of values for airline score

Collected data is plotted in X-Y scatter plot to visualize the pattern for easy identification. Data plots are shown in Fig (1) and Fig(2) for Table(1) and Table(2) respectively.

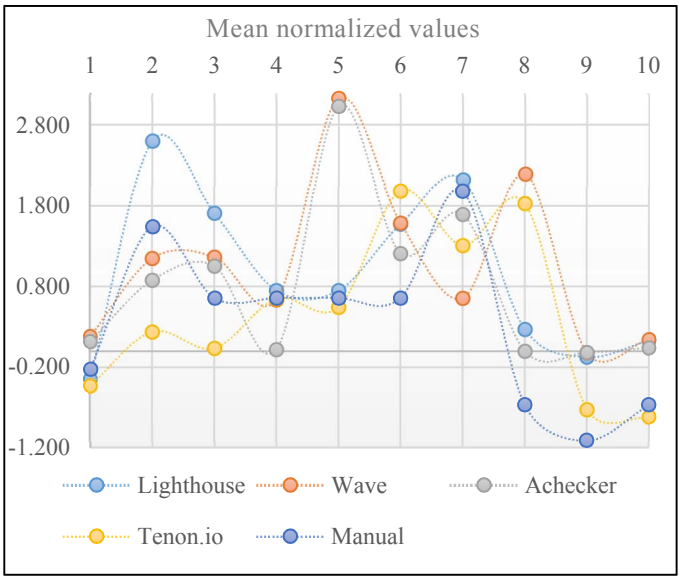


Fig. 2: z-score values plot of values for airline score

Also, we have manually analyzed homepage of all listed websites. The criteria's of WCAG2.1 guideline failing most and number of sites in which those were observed are presented in Table (3).

Table 3: WCAG2.0 guideline violation observed

Issues related to	No of occurrences
1.1.1 Non-text Content	6
1.3.2 Meaningful Sequence	10
1.4.3 Contrast (Minimum)	9
2.1.1 Keyboard	8
2.1.2 No Keyboard Trap	2
2.4.1 Bypass Blocks	3
2.4.3 Focus Order	7
2.4.4 Link Purpose (In Context)	4
2.4.6 Headings and Labels	2
2.4.7 Focus Visible	5
3.2.3 Consistent Navigation	2
3.3.1 Error Identification	4
3.3.2 Labels or Instructions	4

IV. RESULTS

A. Results analysis

Nnormalized data from all methods is plotted for all 10 websites using x-y scatter plot. As shown in figure 1 and figure 2. From the plots it is evident that the websites which has shown less errors using ACheker and WAVE, had performed well with other tools as well as in manual evaluation. That is in accordance with the studies conducted to evaluate associability results using various automation tools.

Analysis of data depicts that American Airlines, Delta and Virgin-Atlantic airways has performed well in all evaluations. Delta website had no error found with WAVE as well as AChecker while Virgin-Atlantic was having only 3 errors in WAVE and 7 errors in ACheker. Score for Lighthouse for these websites was also above 90. Lighthouse is giving more count (200+) for all websites, but comparing the error count with other websites, the score for VAA and Delta is best.

American Airline (AA) has performed best in Lighthouse evaluation with a score of 99. Error score is very less for AA in AChecker and WAVE evaluation tools as well. Tenon has reported slightly higher errors in AA and manual evaluation is also supporting analysis of Tenon with slightly degraded results compared to Delta and VAA.

BA had excellent score with ACheker but all other tools as well as manual evaluation pushed it lower in accessibility score and higher in error count. Similarly, for KLM, AChecker performance was very good with 1 error only while WAVE has highlighted lot of Color contrast issues in KLM homepage, which were found valid in manual evaluation.

Other candidates, Emirates, Qantas, Lufthansa and United Airlines were average with all tools and manual evaluation. However, Cathy pacific is having lot of accessibility issues and need more work to be compliant to WCAG2.0 AA levels.

Visual representation of normalized data shows performance of websites with used methods. Delta has emerged as best performer with all methods closely followed by VAA and American Airlines. British airways have also performed well in most of the scenarios.

The standardized data with z-score shows a major of deviation from mean in terms of standard-deviation i.e. how values are differentiated from mean value in units of standard deviation of series. The plot of z-scores shows a larger range of deviation for Airlines plotted in the middle from number 2 to 8. While plots for 1, 9 and 10 values are in smaller range and coincide with each other in plot. The variation of plot depicts the difference of measures in various evaluation methods. An important result from this study is that for more compliant websites, variation between plots is very less while less compliant websites show a larger variation in values.

B. Conclusion

From the study it is found that among top 10 websites of airline industry, no website has completely passed compliance criteria of WCAG 2.0. Delta Airline has implemented best majors and ranked on top for Accessibility compliance followed by VAA and AA websites. Cathy Pacific need more majors to be taken for improvement of accessibility.

ACheker and WAVE are providing more accurate and quick issue identification in automated accessibility evaluation. Results shown by WAVE and AChecker are also supported and verified using manual evaluation.

For highly compliant websites, all evaluation tools are providing similar results, while variation in results is found more for less compliant websites. That shows the programmatic variation and lack of standardization in evaluation tools.

C. Future work

Detailed study of websites covering internal pages and flows is planned. We will cover more websites from the domain and use statistical analysis tools to evaluate accessibility evaluation tools as well as Airline industry websites.

REFERENCES

[1]WAI, Introduction to Web accessibility, [Online], Available: https://www.w3.org/WAI/.

[2] Gaurav Agrawal, Mayank Singh, Diksha Dani, Devendra Kumar; Evaluating Accessibility and Usability of Airline Websites, July 2019, Advances in Computing and Data Sciences, pp.392-402

[3].Saleh Alwahaishi, Vaclav Snasel, Amine Nehari Talet; Website Evaluation an Empirical study of Arabian Gulf Airlines, February 2010

[4].Ishaq O, A Onuja, H Zubairu; An Analytical Approach to Accessibility and Usability Evaluation of Nigerian Airlines Websites, January 2018

[5].Abid Ismail, K S Kuppusamy; Accessibility of Indian Universities Homepages: An Exploratory Study; Journal of King Saud University - Computer and Information Sciences 30(2) · July 2016

[6] M. K. Baowaly and M. Bhuiyan, "Accessibility Analysis and Evaluation of Bangladesh Government Websites," Proc. of IEEE International Conference on Informatics, Electronics & Vision (ICIEV), pp. 46-51, 2012.

[7] https://www.w3.org/WAI/standards-guidelines/wcag/

[8]https://www.w3.org/TR/UNDERSTANDING-WCAG20/conformance.html

[9]Yakup AKGÜL, Kemal VATANSEVER ; Web Accessibility Evaluation of Government Websites for People with Disabilities in Turkey; Journal of Advanced Management Science Vol. 4, No. 3, May 2016

[10]https://www.deque.com/blog/the-cost-of-accessibility-false-positives/

[11] AChecker, ATRC Web Accessibility Checker [Online], Available: http://achecker.ca/checker/index.php.

[12] WAVE, WAVE Web Accessibility Tool, [Online], Available: http://wave.Webaim.org.

[13] Namkung, Y. Shin, S.Y. and Yang, S;A Ground Theory Approach to Understanding the Website Experiences of Restaurant Customers, Journal of Foodservice Business Research, 2007

[14] O'Connor, P., & Frew, A. J. (2004). An evaluation methodology for hotel electronic channels of distribution. International Journal of Hospitality Management, 23(2), 179–199.Google Scholar

[15] O'Connor, P. (2004). Conflicting Viewpoints on Web Design. Journal of Travel & Tourism Marketing, 17(2/3), 225–230Google Scholar

[16] Wilson, S & Abel, I. (2002), So you want to get involved in E-commerce? Industrial Marketing Management, 31(2), pp. 85–94.Google Scholar

[17] Toms, E. G., & Taves, A. R. (2004). Measuring user perceptions of Web site reputation. Information Processing & Management, 40(2), 291–317.Google Scholar

[18] Xiong L., Cobanoglu C., Cummings P., DeMicco F. (2009) Website Accessibility of U.S. Based Hospitality Websites. In: Höpken W., Gretzel U., Law R. (eds) Information and Communication Technologies in Tourism 2009. Springer, Vienna. https://doi.org/10.1007/978-3-211-93971-0_23

IOT Based Air Discomfort Reliever

Soumi Karmakar,
Fourth Year Student, ECE Dept., IEM.
Kolkata, India
email: soumikarmakar09@gmail.com,

Tapan Kumar Rana
Professor, ECE Dept., IEM.
Kolkata, India
tkrana@rediffmail.com

***Abstract*—Now a days, air pollution is one of the world single biggest environmental risk factors. It is the second leading cause of non- communicable diseases, such as stroke, cancer, and heart disease. In this paper we are presenting an innovative IOT based portable unit for a person while staying outside home. The unit provides fresh air for breathing, relief for sun stroke and remote monitoring. Firstly the unit senses air pollution level and provides pollution free air to breath in a highly polluted area. It also senses environmental temperature and humidity. When the air discomfort level which is an index derived from temperature and humidity, goes high it generates alarm. This will help a person to take necessary preventive action and protect him from health hazard like sun stroke. Heartbeat of that person is also sensed by a heartbeat sensor. Then environmental as well as the personal health data is transmitted through a Wi-Fi module to a useful destination where automatic monitoring and necessary actions are initiated. The proposed unit will find huge application in coming days and give relief to people to a great extent.**

Keywords—Remote monitoring; Sensors; Wi-Fi module; non-communicable diseases; Discomfort index

I. INTRODUCTION

Air Pollution is a very difficult issue to deal with. The sources are chemical, fertilizer, power plants, automobiles, domestic heating and cooking, natural resources like natural dust, volcanoes and sea salts etc. The pollution level is increasing day by day. Air pollution level has crossed the maximum permissible limit and is threatening the existence of human. Approximately, 1.1 billion people across the world breathe unhealthy air. It is responsible for 7 million deaths every year globally. The conventional consumption of fossil fuel has to be eliminated in future to reduce the pollution. Actually lot of research is going on renewable energy sector which is green in nature. Fig1 , taken from Google shows different pollutions.

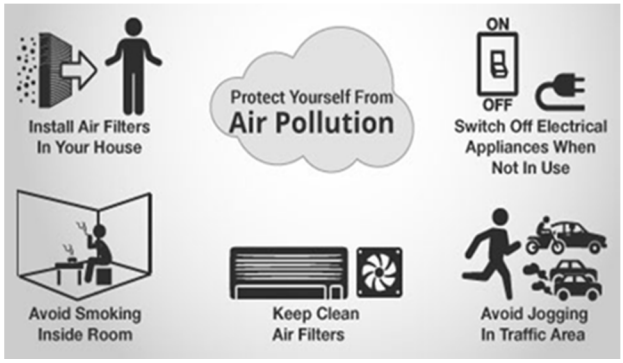


Fig. 1. Protection technique from air pollution

Different technologies are being invented to cope up with this problem. Many papers are available as given in the references list [1] to [10]. In [1] authors presented automatic air purification and adjusting device. The device provides fresh air. In [5] authors introduced a system that controls and detect air pollution in vehicles. The synchronization and execution of the entire system is monitored by a microcontroller. In [6] authors proposed an IOT based low cost pollution control and air quality monitoring system by using Raspberry pi. The system collects data from the sensors, stores the collected data in MySQL database and evaluates by using ThingSpeak open IOT platform. In [9] authors introduced an IOT based smart device to measure the temperature, humidity, carbon monoxide, LPG, particulate matter like PM2.5,PM10 in the atmosphere and monitor the measured data through android applications. So these papers are mainly on air quality controlling and monitoring systems. In our paper, we have designed innovative, useful, low-cost, efficient air purification and monitoring system housed in a portable unit. The unit senses environmental parameters like air pollution level, temperature and humidity. Additionally it is also provided with a human pulse measuring unit. Pollution level, discomfort level and pulse rate are shown in LCD display. When the pollution level is greater than the permissible limit, an air purifier is automatically turned on and provides pollution free air to breath. Nasal mask is used to inhale the fresh air. The unit also gives alarm to the person when the discomfort level goes high so that the person can take necessary action to protect him. Then environmental and health information of that person are sent by Wi-Fi based transmitter to a useful destination where automatic monitoring and preventive actions are taken.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

II. METHODOLOGY

Referring to the fig.2. the pollution sensors are mounted on a bag. The sensed value is processed through a microcontroller. When the level is above danger limit, it operates an electrostatic purifier that emits millions of negative ions into the air forcing dangerous pollutants, allergens, pet dander or odor out of the sphere of clean breathable air. The purified air is tapped through air pipe to the mask. Heartbeat of that person is also sensed by heartbeat sensor. Environmental personal health data is transmitted through Wi-Fi module to cloud. IOT based system monitors and controls from a remote place.

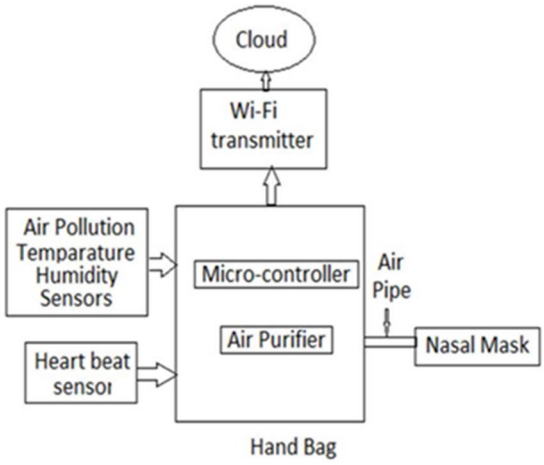


Fig. 2. Block diagram of the hardware

III.DIFFERENT SENSORS AND COMPONENTSUSED

Figures 3 to 6 showing different sensors are taken from Google.

A. Temperature and Humidity Sensor

This sensor senses both moisture and air temperature. Here a basic low-cost digital sensor named DHT11 is used to sense the temperature and humidity of the air.

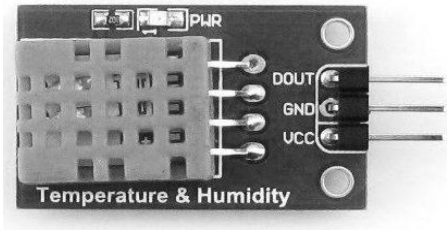


Fig .3. Temperature and Humidity sensor

B. Nasal Mask

Nasal mask has been made, as shown in figure for inhaling pure air generated by the air purifier unit.

Fig. 4. Mask

C. Air Purifier

Fig.5.shows the Air Tamer A310, rechargeable personal or travel air purifier. It emits millions of negative ions and push atomic sized pollutants like viruses, pollen, smoke, molds and dust away from breathing zone. The Air Tamer electrostatic purification radiates a 3-foot orbit of safer air in any direction and creating sphere of protection against harmful pollutants. It has also a break-away connector for extra safety and a conductive lanyard which is adjustable. We can simply wear it around our neck and have a sphere of breathable clear. Its energy efficient technology provides 150+ hours of run time on one charge. In our proposed unit when pollution level is greater than threshold, this air purifier is automatically turned on and the green "on" indicator light starts blinking to indicate that the Air Tamer is cleaning the air.

Fig. 5. Air Tamer

D. ESP8266 module

The ESP8266 is Wi-Fi module with integrated TCP/IP protocol stack that can give access with any microcontroller to Wi-Fi network. In this unit ESP8266 is used to transmit the environmental and health information of a person who is out of home to the cloud by using Thingspeak platform. In this way one can monitor somebody moving outside from a remote place.

It is very much useful to monitor elderly as well as weak persons on road.

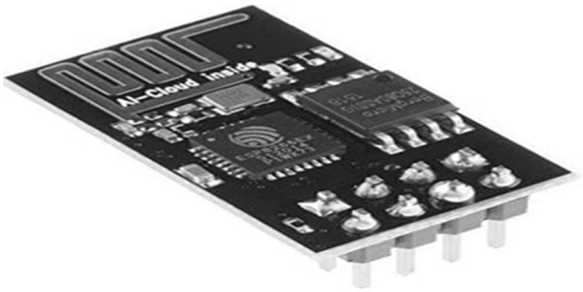


Fig. 6. ESP8266 module

IV. FLOW CHART

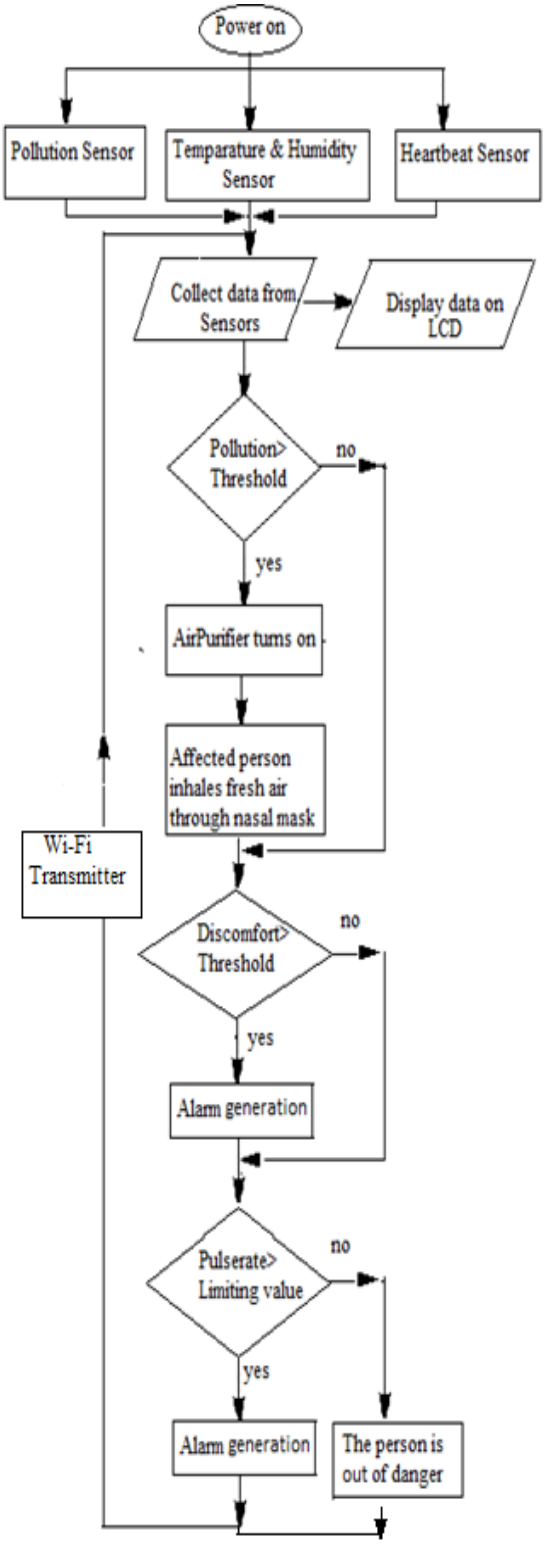


Fig.7. Flow chart

V. RESULTS

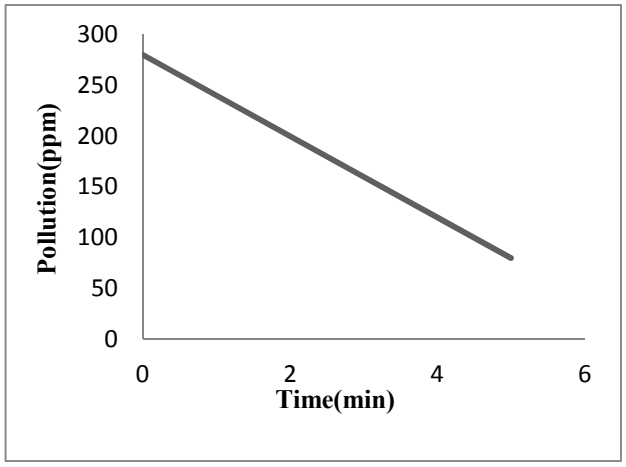


Fig.8. Experimental Graph

Figure 8 is a graphical presentation of change in pollution level after the activation of the air purifier.

Discomfort index DI=T−0.55× (1−0.01H) × (T−14.5), where T is the ambient temperature in centigrade and H is humidity in percent.

A comparison between the theoretical and experimental values of discomfort index is shown in Table1. We have calculated discomfort index by using this formula. When discomfort index value is greater than limiting value, alarm is generated.
The results of the experimental setup with the theoretical calculated values the shown in the following table.

TABLE 1: Discomfort Index

Theoretical Values			Practical Values		
Temperature (°c)	Humidity (%)	DI (°c)	Temperature (°c)	Humidity (%)	DI (°c)
21	50	19.2	22	49	19.9
24	53	21.5	25	50	22.1

IV. CONCLUSION

It is a fact that air quality affects human health. Air-pollution causes many diseases like asthma, stroke, cardiovascular damage, respiratory problem etc. In this paper we have designed a unit for the betterment of mankind. The unit has three parts. IOT technology enhances the importance of the unit and makes it different from other existing solutions. The system can be implemented at a very nominal investment. The unit is portable and everyone can carry it in his bag. Day is coming when this unit will be used as an essential unit for everyone to live. The unit will work in area having Wi-Fi or mobile network connectivity.

ACKNOWLEDGMENT

The authors gratefully acknowledge the management of Institute of Engineering & Management, Kolkata for giving the opportunity and platform for conducting this innovative project work.

REFERENCES

[1]Z. Gai and J. Zhang, "Automatic air purification and adjusting device," 2011 International Conference on Electronics, Communications and Control (ICECC), Ningbo, 2011, pp. 1905-1908, doi: 10.1109/ICECC.2011.6067941.
[2]Cynthia J., Saroja M.N., P. Sultana and J. Senthil, "IoT-Based Real Time Air Pollution Monitoring System", *International Journal of Grid and High Performance Computing*, vol. 11, no. 4, pp. 28-41, 2019. Available: 10.4018/ijghpc.2019100103.
[3] "ESP8266 WiFi Module | Sensors & Modules", Electronicwings.com, 2020. [Online]. Available: https://www.electronicwings.com/sensors-modules/esp8266-wifi-module. [Accessed: 11- Jul- 2020].
[4] Jenifer and D. Aravindhar, "Iot Based Air Pollution Monitoring System Using Esp8266-12 With Google Firebase", *Journal of Physics: Conference Series*, vol. 1362, p. 012072, 2019. Available: 10.1088/1742-6596/1362/1/012072.
[5] S. S. Chandrasekaran, S. Muthukumar and S. Rajendran, "Automated Control System for Air Pollution Detection in Vehicles," 2013 4th International Conference on Intelligent Systems, Modelling and Simulation, Bangkok, 2013, pp. 49-51, doi: 10.1109/ISMS.2013.94.
[6] R.Kiruthika and A. Umamakeswari, "Low cost pollution control and air quality monitoring system using Raspberry Pi for Internet of Things," 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, 2017, pp. 2319-2326, doi: 10.1109/ICECDS.2017.8389867.
[7] "16x2 LCD Pinout Diagram | Interfacing 16x2 LCD with Arduino", *Electronics For You*, 2020. [Online]. Available: https://www.electronicsforu.com/resources/learn-electronics/16x2-lcd-pinout-diagram. [Accessed: 12- Jul- 2020].
[8] G. Sharp, J. Schwartz, D. Buzzeo, E. Desrochers and D. Farrington, "Air Quality Monitoring Systems and Methods", US 7,302,313 B2, 2007.
[9] H. Gupta, D. Bhardwaj, H. Agrawal, V. A. Tikkiwal and A. Kumar, "An IoT Based Air Pollution Monitoring System for Smart Cities," 2019 IEEE International Conference on Sustainable Energy Technologies and Systems (ICSETS), Bhubaneswar, India, 2019, pp. 173-177, doi: 10.1109/ICSETS.2019.8744949.
[10] "The Amazing Benefits of Air Purifiers", *Lifehack*, 2020. [Online]. Available: https://www.lifehack.org/525463/the-amazing-benefits-air-purifiers. [Accessed: 12- Jul- 2020].

Approach for Increasing the Adaptability of Digital Outdoor Advertising

Aleksandr Belov
Department of the Applied Mathematics
National Research University Higher School of Economics
Moscow, Russia
0000-0001-7193-0633

Yaroslav Abramov
Department of the Applied Mathematics
National Research University Higher School of Economics
Moscow, Russia
yaaabramov@edu.hse.ru

***Abstract*— The article describes possible solutions to improve the adaptability of digital outdoor advertising. A mathematical model for placing smart advertisements has been built and an algorithm for finding an effective schedule for displaying advertisements has been proposed. Evaluations of the effectiveness of a set of commercials placed in accordance with the proposed approach are given as a model check. The proposed model can be used to control a distributed outdoor advertising system using IoT technology.**

Keywords*— *DOOH advertisement, IoT, smart advertisement schedule, mathematical model, computer simulation

I. Introduction

A carefully thought-out advertising campaign that stands out from the product can make the business more profitable. Advertising should be targeted to those customers in whom the advertiser is interested. At the same time, "ideal" customers will be people who are interested in viewing ads and are completely satisfied with its content. These requirements can be met by advertising on the Internet, which is probably related to the client, since personalization uses data about the history of visits to sites, search queries, and so on. In this case, an interesting adaptation situation appears on the outdoor advertising market. On the one hand, outdoor advertising indicators differ from online advertising and on the other hand with the development of hyperlocal targeting technologies it becomes possible to collect information about the audience flow for personalizing advertising [1, 2].

Advertisements and ways to convey advertising information to the user are different. They differ in such characteristics as: the cost of placement, the measurability of the final result, the conditions for advertising, etc. The most popular resource of advertising is the Internet. Online advertising uses a wide range of digital opportunities to realize the idea of advertising. Unlike traditional advertising, which is losing its popularity and effectiveness. However, despite the general trend, the volume of outdoor advertising (OOH) is growing steadily and attracts the attention of advertisers [3, 4]. According to Magna Global statistics [5] OOH accounts for 6% of global advertising spending, and the global market share has been growing steadily over the past five years. In some countries, such as France and Russia, it reaches 12% per year. JCDecaux experts are confident that many favourable factors contribute to this type of advertising. Among them [6]:

1. People begin to spend more time outside the home OOH means wide coverage. The audience percentage of this type of advertising varies from 50% to 90% in some markets.
2. Outdoor advertising is a part of the urban landscape that a person sees every day and which cannot be hidden as is happening on the Internet using special utilities.
3. The development of digital technology has given a new look at this format of advertising. DOOH (Digital OOH) or digital outdoor advertising allows you to personalize advertising. Now it has become possible to track the movement of digital people's accounts in real coordinates using technologies based on GPS, Wi-Fi and retina scanners. This helps to accurately determine the position of the advertising consumer and, accordingly, the likelihood of his contact with the advertising message.

A promising direction in DOOH is the use of IoT technology because it allows you to get more information about advertising consumers. Based on the information received you can effectively manage a distributed network of the digital advertising panels [7].

Digital outdoor advertising makes a significant contribution to an optimistic assessment of the prospects of this advertising industry. The market share of digital advertising is increasing every year and the growth of investment in this type of advertising takes the second place among all types of advertising in general.

II. Problem Statement

There is a problem of quantitative calculation of the effectiveness of outdoor advertising. To solve this problem, it is proposed to use methods for evaluating advertising parameters and terminology, which are actively used in online advertising. But the pricing, placement and evaluation of the effectiveness of advertising messages on the Internet and on the street takes place according to different rules. The main criterion for the effectiveness of this type of advertising on the Internet is the clickthrough rate (CTR) [8].

$$CTR = (clicks / impressions) * 100\% \quad (1)$$

This coefficient allows you to show the number of people who accurately noticed the advertisement, as they showed interest in the product and clicked on the link.

To determine the cost of placing an advertising campaign, the CPM (Cost-Per-Mille) and CPC / CPA (Cost-Per-Click / Cost-Per-Action) parameters are usually used. Accordingly,

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

when setting the budget for the CPM, the advertiser pays the number of impressions for the advertisement. In another case, the placement can be carried out in accordance with the CPC / CPA scheme, and payment takes place for the active action that the user performs, for example, a click. Placing commercials on digital screens is a cycle of 6 commercials that replace each other after 6 seconds of playback and 1 second of technical delay. Thus, pricing is formed only on the basis of the duration of the advertising campaign and the geographical location of the advertising screen. Assessing the effectiveness of advertising, you need to pay attention to the number of potential advertising contacts. The assessment of this parameter is based on a statistical analysis of the data set on the main flows (transport, public transport, pedestrians) of advertising consumers located close to each studied advertising object.

In Russia flows of outdoor advertising consumers are estimated using CCTV cameras installed on billboards on transport highways. In addition, pedestrian flows can be estimated according to cellular telephone operators. Generalized characteristics of outdoor advertising consumers, such as: flow rate, distribution of people by age and gender can be obtained using only open sources [9].

Thus, there are several channels for obtaining information about consumers of information, which leads to the need to create a system that could recommend showing advertising regardless of the type and location of the digital advertisement [10]. The goal is to create a decision support system for displaying advertising messages from a common set of advertising campaigns anywhere available for use. It is assumed that the effectiveness of adaptive advertising is measured by the size of the potential audience that is in the field of view of the digital screen.

Consider the task of determining an effective set of commercials for display. Advertising messages will be broadcast on the screen of a street monitor located in an arbitrary geographical location. Viewed advertising messages should be as adaptive (relevant) for people viewing them. In this case, the effectiveness of advertising will be considered only from the point of view of the most significant for the consumer.

III. Mathematical Model of the Digital Outdoor Advertising Campaign

This model will be based on the analysis of statistical data that is collected through open sources for hyper-local audience targeting and special sensors to measure the quantitative characteristics of the flow of people passing through the advertising visibility area (100 m). To achieve greater efficiency, it is proposed to place advertising campaigns without strict reference to geographical location. Other words, the advertising message α_i, which is part of the general portfolio of commercials A can be displayed at any time in any location l_t, $(i = 1 \dots N)$ based on efficiency considerations.

Firstly consider the model for placing advertising messages in one location. Each message has its own set of relevance characteristics. This means that to each advertising message α_i there corresponds a set of its characteristics $\{\gamma_m\}$, $(m = 1 \dots M)$. Then the advertisement portfolio looks like matrix:

$$\begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1(p-1)} & \gamma_{1p} \\ \gamma_{21} & \gamma_{22} & & \gamma_{2(p-1)} & \gamma_{2p} \\ \vdots & & \ddots & \vdots & \\ \gamma_{(n-1)1} & \gamma_{(n-1)2} & \cdots & \gamma_{(n-1)(p-1)} & \gamma_{(n-1)p} \\ \gamma_{n1} & \gamma_{n2} & & \gamma_{n(p-1)} & \gamma_{np} \end{pmatrix}$$

Where γ_{ij}, $i = 1 \dots n$, $j = 1..p$ – vectors of the different dimensions.

For a more convenient operation of this matrix the splitting of the slices of characteristics into "microsections" will be used - all possible combinations of slices. This is necessary to obtain a new advertisement portfolio matrix B with new relevancy ratios for each microsection.

$$B = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1m} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \cdots & \beta_{nm} \end{pmatrix}$$

β_{ij} – the relevance factor j-th to the microsection of the i-th advertising message.

Each location differs from the others not only by location, but also by people who pass through them. We introduce the vector of the characteristic of the location:

$$X_t = \{x_1, x_2, \dots, x_m\}$$

It shows the average number of people from the microsections that passed through the location at i^{th} time.

Then the criterion for choosing advertising video at a given time j is the functional:

$$I_j = \min_{i=1..n} \|a_i - \chi_t\|_2 \qquad (2)$$

To compose an effective, from the point of view of technical and economic indicators, the advertising schedule, it is necessary to minimize the functional $I_{1..T}$ for all time slots, taking into account the restrictions on the number of impressions of advertisements of each type:

$$I_{1..T} = \min_{i=1..n} \|a_i - \chi_t\|_2 \qquad (3)$$

Where T is the final point of time.

The solution of this problem will be sequences of pairs of numbers $(1, j)$ denoting the optimal advertising video at j time. This sequence can be represented as a strongly sparse matrix. For example, for a certain location i, the matrix will look like:

$$M_i = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & & & \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & & & \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \ddots & & & \\ & & & \vdots & & & & \ddots & 0 & 1 \\ & & & \vdots & & & & 0 & 1 & 0 \\ & & & 0 & & & & 0 & 1 & 0 \end{pmatrix}$$

The matrix M_j is called **an effective set of commercials on the i-th location.**

Natural limitations for the proposed model are:

1. Minimum Impressions. Since advertisers can set requirements for the minimum size of an advertising campaign, limiting the minimum number of impressions of an advertising campaign will be natural. It will look like this:

$$\sum_{\text{by columns}} M_i \geq K_j \quad (4)$$

2. Deviation factors. When viewing an advertisement more than a certain number of times, a person has the effect of rejecting the advertised product, which negatively affects the perception of the advertised brand.

$$\sum_{\psi_i \text{ rows, by columns}} M_i \leq \varphi_i \quad (5)$$

IV. Computer Simulation Results

The proposed model was used for computer simulation of the problem of determining an effective set of commercials and a schedule for showing them. Consider the behavior of this system by simulating the display of ads on real data. The computer simulation was carried out on the basis of Microsoft Excel using VBA and Python. The use of this SW is due to the fact that they are widely used among professionals working in the field of advertising. We'll cover 3 geographic locations near public transport stops where we'll place 5 commercials. Analysis of the flows of advertising users in the selected locations was carried out on the basis of data from open sources. We will assume that at any time the aggregate structure of the flow does not change. Data from the Table 1 are showed the following audience profile.

TABLE I. Properties of Locations by Sections

Sections \ # of locations	Location 1	Location 2	Location 3
Male	65,40%	47,40%	36,50%
Female	34,60%	52,60%	63,50%
Young age group	1,24%	4,47%	3,15%
Middle age group	78,00%	72,38%	82,15%
Elder age group	20,76%	23,15%	14,70%
Reach (amount)	5595	10943	1241

This table shows the characteristics of the flow of users of outdoor advertising passing in the specified locations. The figure in each cell is percentage of users belonging to a group with the corresponding characteristic. For example, 65,4% of users in location 1 are men.

The general distribution of people by the clock in the locations is presented at Fig. 1. The abscissa shows the morning time from 7:00 a.m. to 10:00 a.m. The ordinate shows the number of people passing through the specified locations.

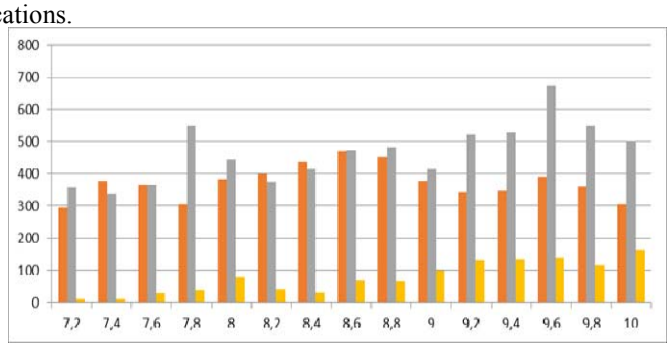


Fig. 1. THE NUMBER OF PEOPLE IN THE LOCATIONS

Legend: location1 – red, location2 – green, location3 – purple.

TABLE II. Characteristics of the Advertising Campaigns by Sections

Sections \ Advertisements	Ad1	Ad2	Ad3	Ad4	Ad5
Male	0,5	0,47	0,2	0,3	0,2
Female	0,5	0,53	0,8	0,7	0,8
Young age group	0,1	0,8	0,05	0,5	0,4
Middle age group	0,2	0,15	0,6	0,1	0,2

By microsections:

$$B = \begin{pmatrix} 0.053 & 0.106 & 0.371 & 0.047 & 0.094 & 0.329 \\ 0.4 & 0.075 & 0.025 & 0.4 & 0.075 & 0.025 \\ 0.01 & 0.12 & 0.07 & 0.04 & 0.48 & 0.28 \\ 0.15 & 0.03 & 0.18 & 0.35 & 0.07 & 0.42 \\ 0.18 & 0.04 & 0.08 & 0.32 & 0.16 & 0.32 \end{pmatrix}$$

Restrictions on the minimum display of advertisements we take K = (50, 30, 60, 40, 70) for ad Ad1, Ad2,..., Ad5, respectively. The coefficients of rejection in the solution of this problem should be set equal for all advertising commercials: φ = 2, and ψ = 5. It means that an advertisement cannot be shown more than 2 times in 5 videos, otherwise it will cause a feeling of rejection of the advertised product.

$$\begin{cases} \sum_{\text{by columns}} M_j \geq K_j, \text{ где } K = (50, 30, 60, 40, 70) \\ \sum_{5 \text{ rows, by columns}} M_j \leq 2 \end{cases} \quad (6)$$

Based on the collected statistics we model the data sets.

After the computer simulations we got the results are presented at Fig. 2 and Fig. 3.

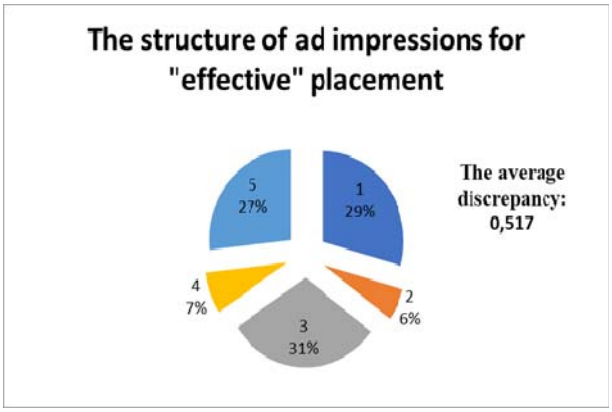


Fig. 2. THE STRUCTURE OF AD IMPRESSIONS FOR "EFFECTIVE" PLACEMENT

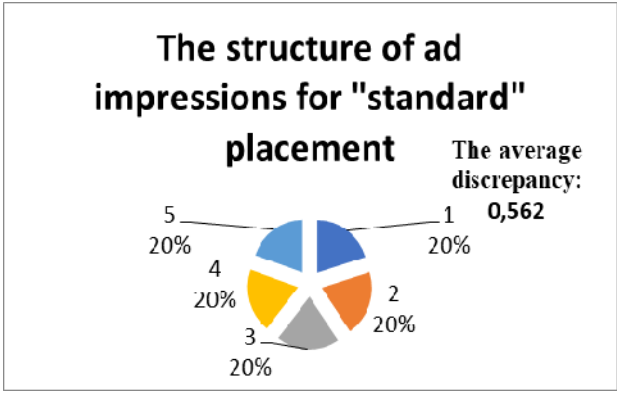


Fig. 3. THE STRUCTURE OF AD IMPRESSIONS FOR "STANDARD" PLACEMENT

Summing up we can say that under the selected initial conditions and assumptions, the discrepancy decreased by 8.57%, and, therefore, the adaptability of the set of commercials increased. This indicator means that the average level of involvement of people who watched commercials in selected places increased by 8.6% solely due to a specially selected schedule. This is a rather large figure in determining the number of promoters in marketing, so the result can be called very good. If you use more advanced sensors to more accurately determine the behavior of human flows by microsegments, this result can be increased up to 10-15%.

V. CONCLUSION

This article proposes a new approach to solving the problems of increasing the effectiveness of digital outdoor advertising.

We can conclude that the developed model contains the basic principles of the decision support system in the display of advertising clips of the adaptive audience. The resulting model can be used as a basis for further research in this field. The development of the SW for the implementation of this model will be based on C ++ / C # to create a friendly interface and optimize data analysis algorithms.

A static model is proposed in the work. To increase the efficiency of constructing a schedule for showing commercials, as well as determining where they are placed on outdoor advertising screens, it is necessary to use a dynamic model that allows you to process data received online. Using IoT technology it is possible to collect information on micro-sections. To do this you need to install a sensor in the selected location that can interact with the gadgets of passing people over Wi-Fi. The device works by the principle of scanning the surrounding space and collects MAC data of devices that are not personal data. Further the system can receive aggregated data and study the dynamics of the flow of people. When creating a dynamic system, it will be necessary to solve the problems of storage and processing of a very large amount of data. To do this, use specialized DBMSs, for example, ClickHouse, developed by Yandex [11].

Using the technologies described in the work, the proposed static model can be transformed into a dynamic one which will allow it to increase its efficiency and the accuracy of forecasting the flow parameters [12]. This means that outdoor advertising will become less toxic to people, reduce the overall level of information noise, and for advertisers to increase levels of awareness and loyalty to the advertised goods and brands.

REFERENCES

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] was the first ..."

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

[1] The Biggest Data Trends for Outdoor Advertising in 2017. Available at https://carto.com/blog/biggest-data-trends-outdoor-advertising-2017/ (accessed 06/02/2020)

[2] K. Li and T. C. Du, "Building a Targeted Mobile Advertising System for Location-Based Services," in Decision Support Systems, Vol. 54, No. 1, 2012, pp. 1-8, doi: 10.1016/j.dss.2012.02.002

[3] G. Chen, J. H. Cox, A. S. Uluagac, and J. A. Copeland, "In-depth survey of digital advertising technologies," IEEE Communications Surveys Tutorials, Volume 18, Issue 3, pp. 2124–2148, thirdquarter 2016, doi:10.1109/COMST.2016.2519912

[4] J. Yan, N. Liu, G. Wang, W. Zhang, Y. Jiang, and Z. Chen, "How much can behavioral targeting help online advertising?" in WWW '09: Proceedings of the 18th international conference on World wide web, April 2009, pp. 261–270, doi: 10.1145/1526709.1526745

[5] Ten Reasons Why Out Of Home Advertising Outperforms. Available at https://magnaglobal.com/ten-reasons-why-out-of-home-advertising-outperforms/ (accessed 02/02/2018)

[6] 10 Reasons Why OOH Advertising Continues Outperform. Available at http://www.jcdecaux.com/blog/10-reasons-why-ooh-advertising-continues-outperform/ (accessed 01/11/ 2018)

[7] Hidayet Aksu, Leonardo Babun, Mauro Conti†, Gabriele Tolomei†, and A. Selcuk Uluagac, Advertising in the IoT Era: Vision and Challenges, 2018, arXiv:1802.04102v1

[8] Osborne A., Coleman R. Outdoor Advertising Recall: A Comparison of Newer Technology and Traditional Billboards, in Journal of Current Issues and Research in Advertising, 2008, vol. 30 no.1, 2008, pp. 13-30, doi: 10.1080/10641734.2008.10505235

[9] Hyperlocal targeting for business with offline points based on Yandex Direct and Google AdWords. Available at http://360-media.ru/Hyperlocal targeting for business with offline points based on Yandex Direct and Google AdWords, (accessed 03/09/2018)

[10] Ayse Nesrin Akoren. Interaction of Outdoor Advertising Improved by Innovative Methods with Digital Art. Procedia - Social and Behavioral Sciences 195 (2015), pp. 799 – 805, doi: 10.1016/j.sbspro.2015.06.178

[11] Programmatic close. Segments Yandex. Audience based on data from external providers. Available at http://360-media.ru/Programmatic close. Segments Yandex. Audience based on data from external providers, (accessed 09/04/2019P)

[12] Belov A. V., Slastnikov S. A. Improving the efficiency of solving discrete optimization problems: The case of VRP//Journal of Physics: Conference Series. 2016. Vol. 692. No.1. p.012050, doi:10.1088/1742-6596/681/1/012050

Permissioned Blockchain-based Security for IIoT

Samira Yeasmin
Department of Computer Engineering
Al Yamamah University
Riyadh, Saudi Arabia
201622195@yu.edu.sa

Adeel Baig
Department of Computer Engineering
Al Yamamah University
Riyadh, Saudi Arabia
a_baig@yu.edu.sa

***Abstract*—Industrial Internet of Things (IIoT) is one of the most trending topics in the present world. It has numerous benefits and its potentiality is unfolding gradually. The implementation of IIoT in mission and safety-critical systems has shown its significance. Since it deals with important and sensitive information, an extensive study is required to address the susceptibility of it towards security issues. Therefore, there have been many proposals to use certificateless signature scheme, machine learning approaches, public key encryption as well as blockchain to improve the security of IIoT. It is important to note that blockchain is playing a significant role in the IIoT technology where the important characteristics of blockchain, immutability, decentralization, tamper-proof, have made a profound impact on the security vulnerabilities of IIoT. Taking this into account, this paper proposes a permissioned blockchain for IIoT that addresses and guarantees to ensure and improve the security vulnerabilities and susceptibility of IIoT towards cyber threats. The permissioned blockchain enabled IIoT ensures a secure medium for device communication, data sharing, and access control. This paper also discusses the security issues of IIoT and presents a comprehensive analysis of some of the proposed blockchain based solutions to improve IIoT security challenges. It addresses the benefits of a permissioned blockchain enabled IIoT over a public blockchain as well as presents a future direction for the upcoming integration of blockchain with IIoT and other industries.**

Keywords— Industrial Internet of Things, IIoT, Blockchain, Permissioned Blockchain, IIoT security issues, security and privacy, cyber-attacks.

I. INTRODUCTION

Internet of Things (IoT), an emerging technology, is being implemented all over the world and is providing solutions to emergent problems. Industrial Internet of Things (IIoT) is a subset of IoT which connects industrial assets and provides an optimal solution to the arising difficulties of the industries. It focuses on the industrial domain and produces data that helps to make important decisions. It has the potentiality to improve and enhance industrial assets such as machines, control systems, business processes, computers, etc. that enable intelligent industrial operation. Networked smart power, manufacturing, medical, and transportation is some of the use cases of IIoT [1]. IIoT is having a great impact on the economy and it is believed that it will have a significant influence on the mission and safety-critical systems.

Although IIoT seems very similar to IoT, there remain many differences between them in terms of use cases [2] and security requirements that have been studied in [3]. The major difference between these two is how they are used and what kind of service they provide. Consumer-level devices mainly provide their service through IoT. Whereas, the most important area where IIoT is used is in the mission-critical systems. If IoT and IIoT are further compared, a number of dissimilarities can be found in their respective architectures, communication, connectivity, data volume, latency, speed, and scalability.

IIoT helps in making improved business decisions and provides better scalability, connectivity, and efficiency. However, the centralized nature of IIoT devices makes it vulnerable to many cyber threats. The overall security issue of IIoT gets aggravated because of the heterogeneous networks used by different devices. Since IIoT devices are being used in mission and safety-critical systems, security vulnerabilities need to be addressed and improved.

To address the emerging security challenges of IIoT, there have been many proposals to use a blockchain because of its feature of decentralization. The integration of blockchain with IIoT has numerous benefits in improving and solving some of the security issues. A blockchain is a Distributed Ledger Technology (DLT), cryptographically-sealed and temper-resistance, connecting multiple computers in a peer-to-peer (P2P) network that works without any third-party involvement. Using consensus algorithms, such as Proof-of-Work (PoW), Proof-of-Stake (PoS), Proof-of-Authority (PoA), Proof-of-Elapsed-Time (PoET), Byzantine Fault Tolerance (BFT), Practical Byzantine Fault Tolerance (PBFT), etc., new blocks are added to the network [3]. Although the primary application of blockchain was to store financial data in a secured fashion, there have been numerous applications of it.

The rest of the paper is organized as follows: Section II focuses on the background study of IoT and IIoT. Section III analyzes the security issues of IIoT. In section IV, some proposed solutions by different authors to improve the security of IIoT have been discussed. Section V proposes a Permissioned Blockchain enabled IIoT to improve and address the security challenges of IIoT. This section also discusses how a permissioned blockchain would guarantee the security of an IIoT environment. Section VI and VII present a future direction and conclude the paper.

II. BACKGROUND

An increasing number of everyday objects is getting connected to the internet. These objects have computing capabilities and transfer data over a network that does not require any human interaction. They are equipped with sensors and actuators [4] that enables the communication between them so that important and useful decisions can be made. This network of smart objects is known as the Internet of Things (IoT). The main concept of IoT is the connection between things

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

or objects, such as Radio Frequency Identification (RFID), sensors, actuators, mobile phones, etc., which interact with each other to accomplish a particular goal [5]. It connects the physical objects either through a wired or wireless network [6]. It consists of smart devices that have a self-processing unit along with having the capability to exchange information. These smart devices help in controlling the overall network. IoT enables numerous and novel applications, especially in the industrial sector.

The traditional industrial sector consists of numerous independently operating production systems. They have little or no communication among them [5]. An emerging proposal is to interconnect them so that important and useful decisions can be made while making the production system more efficient. The result of this emerging trend is the Industrial Internet of Things (IIoT). It integrates the conventional industrial infrastructure with the internet that enables the industrial sector to be highly flexible, faster, more efficient, and enhances the production system as well as optimize business processes.

Being a subset of IoT, the Industrial Internet of Things (IIoT) interconnects industrial systems and physical objects using sensors and actuators. IIoT integrates wireless sensor networks, communication protocol, and internet infrastructure to connect devices and provide intelligent and efficient industrial operations [7]. It is a technology that deals with machine-to-machine communication which consists of various network-connected industrial machines. They generate, process, and exchange a huge amount of data that is related to the mission and safety-critical infrastructures. They also share confidential information using this communication protocol and thus are subjected to many cyber-threats. For the correct operation of the machines in the IIoT, ensuring the integrity of data is very important. IIoT is an important element for the future of the industry. It is the transformation of IoT from commercial level to industrial level [6]. Some of the IIoT applications include connected medical devices for patient safety, Large□Scale SCADA Control, autonomy, and many more [1]. The IIoT is based on IoT that provides seamless communication between devices to increase productivity, and help the business make more intelligent and important decisions [4].

With the rapid development of IIoT, there have been some issues that need to be resolved. Energy efficiency, real-time performance, coexistence, interoperability, security, and privacy are to name some. One of the major requirements that IIoT needs to address is security. Security and privacy have always been a concern for IoT devices, the same goes for IIoT. Since IIoT is an open, distributed, and heterogeneous system, implementing security becomes a great challenge. With the help of IIoT, industrial infrastructure becomes easily vulnerable to cyber threats and attacks. The reason behind this could be the industrial infrastructure becoming more connected and intelligent by sharing information with the cloud and helping the industries achieve efficiency and enhancement in its performance. IIoT devices make decisions transparently [4] and share them among the stakeholders to help with asset tracking. Transparency of sensitive data attracts malicious attackers to take control of the network. Moreover, IIoT deals with the mission and safety-critical systems that generate a huge amount of data helping in taking intelligent and important industry-related decisions. Thus, securing the generated data and communication protocol becomes extremely important. IIoT devices are susceptible to Man-in-the-Middle (MITM), Device hijacking, Distributed Denial of Service (DDoS), Permanent Denial of Service (PDoS) [8]. These security issues can have a drastic impact on the industrial infrastructures implementing IIoT since they generate, process, and exchange valuable and sensitive data. Therefore, there is a crucial need to improve the security of IIoT in order to protect mission-critical systems and industrial infrastructures.

III. IIoT Security Issues

IIoT deals with an enormous amount of data which helps in taking substantial decisions for various industries that are implementing IIoT in their fields. Hence, it becomes extremely important to address the IIoT security issues.

A. Insecure IoT gateways

IoT gateways, used in IIoT systems, are responsible to connect IIoT devices with the service cloud [9]. IoT gateways help to access data from sensors and analyze these data to make critical decisions for the industry. Since the gateways are the medium for communication between the IIoT devices, a secure design is required. IIoT devices become vulnerable to cyber threats such as MITM and DDoS. If IoT gateways are not secured enough and get compromised by these cyber threats, the whole IIoT network will be at serious risk. That is why it becomes very important to take necessary measures to enhance the IoT gateways.

B. Inefficient and insecure MQTT protocol, server, and access control mechanism

The IIoT system architecture consists of three layers [10]. Each layer is responsible for its respective tasks. The layers are the following.

1) Data acquisition layer: The main responsibility of this layer is to control the devices of IIoT. The sensors collect data that provide information to take the necessary steps in that specific situation. According to [10], there remain some security issues on this layer such as an IoT botnet virus. It exploits the vulnerability of the IoT nodes by building huge botnet attacks and launching DDoS attacks.

2) Data transmission layer: The second layer is responsible to combine the sensor network, mobile networks, and the internet. It requires a communication medium to combine the above mentioned networks. The communication protocol usually used by IIoT is MQTT which does not provide any encryption and authentication [10]. Data is transmitted in plain text. Therefore, it becomes subjected to security threats since intercepting data while communicating with other devices becomes easier.

3) Data processing layer: This layer requires a communication server, a historical data server, and remote monitoring terminals. An MQTT proxy server is used with a simple firewall and access control mechanism [10]. The firewall is so simple that it cannot identify the forged packets that are conforming to the protocol and access control rules. So, they bypass the firewall very easily.

Therefore, for secure data communication, a secured communication protocol is required. Again, data sharing becomes vulnerable to cyber-attacks since data transmits in plain text using MQTT. Moreover, there should be an access control protocol so that data is accessed only by authorized parties. Consequently, confidentiality, integrity, and availability of data will be achieved.

C. Insecure Cyber-Physical System (CPS)

A Cyber-Physical System (CPS) combines computational resources with physical components which include autonomous unmanned aerial vehicles (UAVs), self-driving cars, and home automation systems [11]. According to [5], CPS is the foundation IIoT since it deals with handling a huge amount of information to better control and monitor processes [11]. Sensors are used to receive information and actuators control the physical processes of CPS [12]. It is composed of electronics and monitors everything through sensors. It uses IIoT for its better mechanism and connectivity. Since CPS is now using IT infrastructure for its operation and interconnectivity is a part of its process, it has susceptibility towards many security vulnerabilities which also becomes a security issue for IIoT. The CPS requires extensive use of a secured medium of communication as control commands and feedback loop signals usually use an open network to transport [11] which makes it susceptible to various security threats. The CPS is prone to trojans, viruses, Denial of Service (DoS), and DDoS attacks. According to [5], CPS is vulnerable to cyberattacks in terms of the following.

1) Electronics: are vulnerable to physical attacks, such as hardware, side-channel attacks, and reverse engineering.

2) Software: is susceptible to malicious attacks including trojans and viruses.

3) Communication protocol: Protocol attacks, for instance, MITM and DDoS attack.

4) Human operating Cyber-Physical Production System (CPPS): Social attacks, as for example phishing and social engineering.

D. Sudden integration of Information Technology (IT) and Operation Technology (OT)

The task of Industrial Control Systems (ICS) is to deal with critical infrastructures. They help in controlling and monitoring them [9]. According to [13] IIoT is a combination of Information Technology (IT) and Operation Technology (OT). Earlier, the ICS was not using IT infrastructure, but with the introduction of IIoT, it has also become interconnected while doing its job. With the usage of IIoT, the traditional role of ICS has been redefined and has greatly impacted the industries. However, the connectivity to the internet and the sudden integration of IT in ICS has made it vulnerable to cybersecurity risks. Since OT systems were not designed keeping security in mind, it is increasing the risk of cyberattacks for IIoT. According to the State of Industrial Cybersecurity 2018 report [14], most of the companies (65%) implementing IIoT believe that there is a high risk of security vulnerabilities and attacks with the sudden integration of OT in the IIoT environment. Unprotected network connection of the OT environment as well as unidentified security issues because of combining IT and OT is risky for the IIoT technologies. Therefore, secure IT and OT as well as secure ICS is required to save IIoT from getting compromised.

E. Insecure SCADA

SCADA is a network of intelligent devices that control and monitor some specific machines or systems or processes [12]. The machines or systems are connected to intelligent devices with the help of sensors. It helps in studying the output generated by the devices and make important decisions. SCADA provides control over the software layer with Programmable Logic Controllers (PLC) that are a part of IIoT [15]. This makes SCADA a segment of IIoT and that is why it is an important industry where IIoT is being implemented. According to [16], SCADA systems have security vulnerabilities that also become a threat to IIoT. These security issues include social engineering, trojan, worm, DoS attack, etc. Therefore, secured SCADA is required to make IIoT efficient and enhance its functionalities.

IV. Proposed Solutions for IIoT Security Using Blockchain

Table I highlights the proposed solutions by several authors to resolve some of the security issues of IIoT. The table shows that the authors in [17], proposes a framework for IIoT that uses blockchain technology so that data can be collected while saving energy as much as possible, and data sharing in a secured manner can be achieved. The framework is designed on a public blockchain, Ethereum, to maintain a tamper-proof shared ledger as well as eliminate the need for an intermediary. It is simulated against the famous blockchain security issues and attacks so that the security goals can be accomplished. The attacks and issues include smart contract vulnerabilities, eclipse, and majority attacks. In [18], a blockchain system is proposed that uses the PoW credit-based algorithm to design secure IIoT devices that will use less power. Since IoT devices used in industries consume a lot of power that becomes an overhead for the industries, therefore, a DAG-structure blockchain is used in [18] that will significantly reduce the power consumption as well as guarantee a scalable, secure blockchain system for IIoT. The system performs well since it is designed to handle the Double Spending and Sybil attacks of blockchain.

A Fabric blockchain is proposed to be used for secure data transmission in [19]. Transaction certificates are used to secure the transmission of data and newly added blocks are authenticated to ensure security. Without the certificates, the transactions remain invalid. Data is remained to be unaffected if a single node is tampered because of using a hash algorithm. Data transaction security is ensured with the use of public/private key pair. Authors in [20], used Ethereum blockchain powered by Attribute-based encryption (ABE) to provide fine-grained access control in IIoT. The main goal of the proposal is to secure the data sharing scheme of the supply chain industry using blockchain. Nodes need to be registered on the blockchain according to their roles so that access control mechanisms can be defined in the smart contract. The paper introduces a secure data sharing scheme where a single point of failure attack is discussed and the PoW consensus algorithm is used.

TABLE I. PROPOSED BLOCKCHAIN SOLUTIONS TO SECURE IIOT

Year	Authors	Proposed Solution	IIoT Security Problem Addressed	Blockchain Approach
2019	Liu et al. [17]	A blockchain enabled energy efficient and secure data collection and sharing scheme	Secure data sharing	Ethereum
	Huang et al. [18]	PoW credit based consensus algorithm	Efficient access control scheme	DAG-structured blockchain
	Liang et al. [19]	A dynamic secret sharing mechanism in data transmission technique using power blockchain	Secure data transmission	Fabric
	Wen et al. [20]	Secure data sharing by combining supply chain, blockchain, and IIoT	Secure data sharing	Ethereum
	Wan et al. [21]	Improve processing power, security, and privacy of IIoT	Secure data sharing and communication	Private
2020	Wang et al. [22]	Reputation scheme to certify the miners are trustworthy and make the communication secure	Trust scheme for IIoT	Ethereum
	Lu et al. [23]	Secure data sharing architecture using blockchain	Secure data sharing	Permissioned
	Shen et al. [24]	A secure device authentication mechanism	Secure device and data communication	Consortium

A lightweight and easily expandable architecture using private blockchain is designed for smart factories, eventually IIoT, in [21]. A private blockchain is used to ensure security and privacy. Moreover, the architecture consisted of five layers, each layer is responsible to secure itself. Verification mechanism, time-stamped, whitelist, and blacklist mechanisms are used and introduced in the architecture so that smart factories implementing this architecture can prevent malicious traffic and erroneous data injection. Moreover, whitelist and blacklist mechanisms, dynamic verification, asymmetric encryption, and hash have been used to address security issues. A reputation scheme is proposed by the authors in [22] where a trust-based consensus mechanism named PoRX is used and can be applied on top of the existing PoX consensus algorithm of blockchain. The scheme helped in certifying miners of the blockchain to secure the communication medium. The reputation of a node remains the same or gets increased if it can successfully add a block to the chain by getting a confirmation from the consensus nodes. On the other hand, the reputation will get decreased if the nodes cannot produce blocks within a certain time limit and fail to achieve the consensus. In this way, nodes with a decreased reputation will not be able to participate and will be considered abnormal. A trust-based communication will be achieved since only reputed nodes will participate in the chain network.

A permissioned blockchain has been proposed to use for secure data sharing in IIoT by the authors in [23]. Designed for multiple parties, the actual data is stored on local providers where data is retrieved only upon request. This mechanism ensures secure data sharing because only permissioned parties are able to access the data since data is provided if requested by registered participants of the blockchain. In [24], the authors propose to authenticate IIoT devices by using a consortium blockchain along with an extended version of Identity-based signature (IBS) where there is no requirement for a public key certificate. Since it uses consortium blockchain, there is no third party involvement in the authentication process. Moreover, the IBS reduces the overhead of issuing digital certificates. Instead of using the public key, the process is turned into a blockchain writing process to ensure device authentication along with secure data communication.

Therefore, from the above discussion, blockchain is contributing significantly to improve the security of IIoT in terms of (a) secure data transmission, sharing, and communication (b) efficient access control mechanism and (c) trusted communication scheme. Thus, this paper explores how permissioned blockchain can be integrated to improve the IIoT security and how the unique characteristics can be useful in this regard. A permissioned blockchain only allows the participation of nodes only if they are permitted, authorized, and authenticated. For this purpose, the next section discusses permissioned blockchain and its implications to enhance and improve IIoT security.

V. PERMISSIONED BLOCKCHAIN ENABLED IIOT

With the rapid growth of IIoT, a huge amount of data is being generated which might become a bottleneck in meeting the Quality-of-Service (QoS) of this technology. Again, the security requirements of IIoT is also getting increased with the development and involvement of this technology in different industries. The importance of improving the security of IIoT is an emerging issue to solve. Blockchain technology, therefore, can be helpful in improving the security problems with its unique characteristics of storing transactions in a secured fashion and decentralization ability. It can be useful in processing the massive amount of data generated by IIoT. Moreover, this approach will further help in reducing a single point of failure because of its decentralized nature. The cryptographically secured network and immutability nature of blockchain will ensure data security and privacy. The distributed and decentralized attribute of blockchain makes it attractive to replace the existing and conventional solutions for the security issues of IIoT.

A blockchain is a DLT that is cryptographically-sealed and temper resistant. It is a connection of multiple computers in a

peer-to-peer (P2P) network that works without any third-party involvement. There have been numerous implementations of blockchain that have unveiled many unprecedented meanings of this technology [3]. There are three types of blockchains namely public, consortium/permissioned, and private and they differ in how they are implemented in different industries [3]. Despite having numerous benefits over a traditional database, one notable drawback of this technology is the scalability problem [3]. The growing number of devices might become an obstacle in the upcoming applications of blockchain. Furthermore, the scalability problem leads to a costly and slow chain network. However, while the scalability issue promotes low performance, a permissioned blockchain enables a closed network and provides a highly scalable, cost-effective, and faster chain network.

This paper proposes to use a permissioned blockchain to improve the security of IIoT because of its unique features such as identification and authentication of participants, restricting their access and performance of transactions across the chain network.

A. Proposed Permissioned Blockchain for IIoT

A permissioned blockchain consists of several organizations managed together by blockchain and these organizations run one or more nodes. The blockchain is called a permissioned blockchain because only allowed and permitted nodes or participants can contribute to the network and one or more nodes work together to control and restrict access of members of the chain network. Each organization has members of its own and allows them to participate only if they are authorized. The organizations define specific roles to the members so that they can perform transactions in the network. Two or more organizations jointly form the blockchain and record transaction data. A permissioned blockchain has advantages such as decentralization, immutability, cryptographically-sealed similar to a public blockchain [3]. However, it is specifically used to ensure the security of an organization by making it a closed network and authenticating the participants' identities, also restricting their access. A permissioned blockchain can be well suited and applied to the IIoT network. Permissioned blockchains provide a faster, more secure, and privacy-oriented network which is a need for IIoT machine communication.

Therefore, a permissioned blockchain is appropriate to ensure the security of IIoT that will allow only identified and authenticated members to participate in the network and will permit device communication only between the authorized devices. All the other devices outside the network are considered malicious to make the network secure from the cyberattacks of communication protocol. One of the components of the blockchain is a Certificate Authority (CA) that generates digital certificates for the members of the organization to guarantee secure communication across the network. The CA generates key-pairs of public and private keys which helps in verifying identities of the members of the network. Moreover, a smart contract defines the roles of the members after the CA generates keys and guarantees the security of the blockchain. The smart contract helps in assuring that only allowed and permitted members are participating in the network and performing transactions. The transactions are events as data access, storage, sharing, and monitoring. When a new entity performs a transaction, only a limited number of consensus nodes are authorized to verify the transaction and generate a new block after the verification process. The consensus algorithm used in this proposal is Practical Byzantine Fault Tolerance (PBFT). After the verification, the new block is added to the network and all the other members of the organizations are notified so that they can update their ledger.

B. How it works

The permissioned blockchain consists of a CA component that works similar to a traditional CA but within the blockchain so that it remains secured with the help of the attributes of a permissioned blockchain and does not get susceptible to the security issues of the traditional CA. The CA generates certificates to the identities of the organizational members of the blockchain. It generates two types of keys, public and private keys that are used while performing transactions. The identity fills the required information to generate the certificate and a key pair. The key pairs and certificates are stored on the blockchain securely as well as on the smart contract of the chain network. The status of the certificates is stored on the smart contract for further processing.

After generating certificates for the identities, the smart contract defines access control mechanisms to the different identities based on the stored information of the certificates. The identities get access control mechanism defining their roles. Hence, the role-based access control mechanism ensures identities perform transactions according to their roles, and transactions are executed only if their roles provide the permission to do so.

When a member of the organizations invokes a transaction, it is signed with their public key generated by the CA. A limited number of consensus nodes ensure the transaction is according to its role, authenticate it, and guarantee if the public key attached to the transaction is the same as the CA. The consensus nodes, following the PBFT consensus mechanism, approves or disapproves the transaction based on the public key and access control mechanism. Moreover, the consensus nodes maintain a State Merkle Tree (SMT) [25]. This tree stores and verifies the state of all entities, codes, and data of the smart contract so that they can verify the access control mechanism. The SMT gets updated if the transactions are executed and its Merkel root becomes a part of the block header. The storage of smart contract in the consensus nodes allow the tracking of the certificates of the member who performed a transaction. The Merkle root also verifies the authenticity of a block by using the public key of the member.

After executing the transaction, all the members of the network get notified about the addition of the new block to the network and the organizations also update their ledgers accordingly to achieve consistency. If a transaction gets disapproved by the consensus nodes, it will not be executed but the information will be stored on the blockchain although the state of the ledger will not be updated. The disapproved transaction information will help in the future to figure out what happened across the network and who wanted to invoke a transaction but was not allowed to do so. This adds an extra layer

of security to the network and help in tracking transaction information.

C. Discussion

IIoT connects different industrial devices that work together and generate a large amount of data including valuable and sensitive information. It collects and analyzes this data in order to deliver new insights. Therefore, it is extremely important to secure the communication medium between the IIoT devices and protect the valuable data, since compromising the IIoT network will have a serious consequence and will affect the industry and its infrastructures heavily. Using a permissioned blockchain that allows only selected and authorized participants to join the network and perform transactions, will certainly help in improving the security of IIoT. One of the advantages of a permissioned blockchain is having the identity management capability that guarantees security and transaction execution based on the participants' identities. Identification of the participants will ensure the valuable and sensitive data is accessible to only authorized parties of the industry and will protect the network from a MITM attack. The chain data in a permissioned blockchain is shared only among a selected number of participants that ensures the security of the network. Moreover, it uses more efficient consensus algorithms as well as allows us to use more than one algorithm to secure the network as much as possible. The use of algorithms depends on the organization's needs. Since a permissioned blockchain guarantees a closed network, the performance of this network is more efficient and enhanced than a public blockchain because of the participation of only a fewer number of authorized participants. On the contrary, a public blockchain allows anyone to join the network and perform transactions as well as share and see the chain data. Hence, a permissioned blockchain is more secure than a public blockchain because of having identity management, allowing to use more than one consensus algorithms, restricted and identity-based transactions, and better performance which helps in improving IIoT security.

The proposed permissioned blockchain enabled IIoT ensures to improve and enhance the security of IIoT in terms of secured device communication, server, data sharing, and access control mechanism. After only allowed participants join the network, they have restricted access to the network, and an access control mechanism ensures a limited execution of transactions. A CA component of this proposed permissioned blockchain enables the network to be secured from unauthorized access and execution of transactions. A permissioned blockchain is well suited for a secured IIoT network of devices because an organization can define the roles of the participants using smart contracts and restrict their access to the network. The roles also define which member can read the data of the network and execute some write operations. Consensus nodes play a vital role across the network. It ensures restricted execution of transactions and improves the security of IIoT device communication. To achieve privacy and security, this implementation is desirable since it defines the roles and access control mechanism of the participants, uses a smart contract that stores certificate details to guarantee controlled transaction invocation, consensus nodes verify and authenticate the transactions as well as stores the state of all entities, codes, and data of the smart contract. Moreover, in a permissioned chain network, transactions are not anonymous between the participants which gives the organization of the blockchain an opportunity to check who performed transactions and it helps the admin to ensure security and privacy. Thus, transparency and trust between the authenticated parties are established which is an essential requirement to enhance the security of IIoT. A permissioned blockchain enabled IIoT is suitable to improve the security of the whole network and ensure privacy while communicating with other devices.

Table II shows a comparison between permissioned and public blockchain. Unlike a public blockchain, participants' identification and verification as well as limitation and restriction of user access to the network are ensured by a permissioned blockchain. It also guarantees the visibility and transparency of the transaction execution across the chain network which is a security requirement of IIoT. Moreover, for security reasons, transaction validation is done by a limited number of validators on a permissioned blockchain. On the contrary, a public blockchain can be joined by any user and it does not restrict users' access and execution of transactions. The table also presents different platforms available to implement permissioned and public blockchain.

VI. Future Work

An extensive study shows that blockchain is useful when integrated with IIoT. It is expected that blockchain will be revolutionary for IIoT technology. In future research, the implementation of the proposed Permissioned Blockchain enabled IIoT will be performed. It will also be evaluated to analyze the security and performance of the permissioned blockchain. Future work will include a more extensive study of the proposed blockchain along with evaluating the impact of the design choices in terms of security and privacy. Ongoing work is being performed to analyze the effect of a permissioned blockchain, such as Hyperledger Fabric, on securing IIoT communication. Moreover, it is important to discuss the security issues of blockchain that might affect IIoT systems as well. Since blockchain comes in three different types namely public, permissioned/consortium, and private, it is important to study the behavior of each one of them. A future study is also required to learn how the security vulnerabilities of blockchain are affected or enhanced with the use of IIoT. Since both of the technologies are emerging and unfolding its significance, it will

TABLE II. Permissioned and Public Blockchain Comparison

Features	Permissioned Blockchain	Public Blockchain
Identity Management	Yes	No
Limited Action	Yes	No
Anonymity	No	Yes
Restricted Transactions	Yes	No
Transaction Validation [26]	Limited	Open
Consensus Algorithms	PoS, BFT, PBFT	PoW, PoA, PoS, PoET
Platforms [3]	Hyperledger Fabric, Quorum, Corda	Bitcoin, Ethereum, Litecoin

be exciting to discover what unprecedented meaning it is holding for us in the future and how IIoT will affect blockchain security issues. The need for such future work is required where a discussion will be made to understand how the security vulnerabilities of blockchain are going to affect different industries and what is to be expected in the future.

VII. Conclusion

Blockchain, being one of the most emerging technologies, has shown its potentiality in many industries, which includes IIoT as well. The integration of blockchain with IIoT has shown significant revolution and numerous benefits. The emergence of IIoT is going to have an important impact on different industries. Therefore, it is essential to learn and address the security vulnerabilities of IIoT. In this paper, a comprehensive study has been done to find the security issues of IIoT. There have been many proposals to use blockchain in IIoT to resolve cyber threats and attacks. This paper analyzes the main security challenges of IIoT that must be addressed as well as discusses existing proposals to solve IIoT security threats. Moreover, a Permissioned Blockchain enabled IIoT has been proposed to address the security vulnerabilities of IIoT in terms of secured device communication, server, data sharing, and access control mechanism. The use of a CA component, smart contract, and consensus nodes, as well as the performance of restricted transactions in a permissioned blockchain ensures privacy and security across the network. It also studies the importance of a Permissioned Blockchain enabled IIoT over a public one that ensures and guarantees the security of the network. Therefore, the DLT of the permissioned blockchain is going to address the susceptibility of IIoT towards cyber-attacks and is appropriate for the IIoT environment that ensures a faster, more secure, and more privacy oriented network.

References

[1] Hwaiyu Geng, "THE INDUSTRIAL INTERNET OF THINGS (IIoT)," in *Internet of Things and Data Analytics Handbook,* Wiley, 2017, pp.41-81, doi: 10.1002/9781119173601.ch3.

[2] E. Sisinni, A. Saifullah, S. Han, U. Jennehag and M. Gidlund, "Industrial Internet of Things: Challenges, Opportunities, and Directions," in *IEEE Transactions on Industrial Informatics,* vol. 14, no. 11, pp. 4724-4734, Nov. 2018.

[3] S. Yeasmin and A. Baig, "Unblocking the Potential of Blockchain," *2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA)*, Ras Al Khaimah, United Arab Emirates, 2019, pp. 1-5, doi: 10.1109/ICECTA48151.2019.8959713.

[4] Wang, Q., Zhu, X., Ni, Y., Gu, L., & Zhu, H, "Blockchain for the IoT and industrial IoT: A review", *Internet of Things*, 100081, 2019.

[5] A. Sadeghi, C. Wachsmann and M. Waidner, "Security and privacy challenges in industrial Internet of Things," *2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC)*, San Francisco, CA, 2015, pp. 1-6, doi: 10.1145/2744769.2747942.

[6] A. Karmakar, N. Dey, T. Baral, M. Chowdhury and M. Rehan, "Industrial Internet of Things: A Review," *2019 International Conference on Opto-Electronics and Applied Optics (Optronix)*, Kolkata, India, 2019, pp. 1-6, doi: 10.1109/OPTRONIX.2019.8862436.

[7] T. Alladi, V. Chamola, R. M. Parizi and K. R. Choo, "Blockchain Applications for Industry 4.0 and Industrial IoT: A Review," in *IEEE Access*, vol. 7, pp. 176935-176951, 2019, doi: 10.1109/ACCESS.2019.2956748.

[8] Rambus, "Industrial IoT: Threats and Countermeasures," Rambus. [Online]. Available: https://www.rambus.com/iot/industrial-iot/. [Accessed on 21-Mar-2020].

[9] X. Yu and H. Guo, "A Survey on IIoT Security," *2019 IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS)*, Singapore, 2019, pp. 1-5, doi: 10.1109/VTS-APWCS.2019.8851679.

[10] H. Chen, M. Hu, H. Yan and P. Yu, "Research on Industrial Internet of Things Security Architecture and Protection Strategy," *2019 International Conference on Virtual Reality and Intelligent Systems (ICVRIS)*, Jishou, China, 2019, pp. 365-368, doi: 10.1109/ICVRIS.2019.00095.

[11] M. Atif et al., "Soft Computing Techniques for Dependable Cyber-Physical Systems," in *IEEE Access*, vol. 7, pp. 72030-72049, 2019, doi: 10.1109/ACCESS.2019.2920317.

[12] H. Boyes, B. Hallaq, J. Cunningham, and T. Watson, "The industrial internet of things (IIoT): An analysis framework," *Computers in Industry*, vol. 101, pp. 1–12, Oct. 2018.

[13] A. C. Panchal, V. M. Khadse and P. N. Mahalle, "Security Issues in IIoT: A Comprehensive Survey of Attacks on IIoT and Its Countermeasures," *2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN)*, Lonavala, India, 2018, pp. 124-130, doi: 10.1109/GCWCN.2018.8668630.

[14] E. Bajramovic, D. Gupta, Y. Guo, K. Waedt, and A. Bajramovic, "Security Challenges and Best Practices for IIoT," In *INFORMATIK 2019: 50 Jahre Gesellschaft für Informatik–Informatik für Gesellschaft (Workshop-Beiträge)*, Gesellschaft für Informatik eV, 2019.

[15] G. Falco, C. Caldera and H. Shrobe, "IIoT Cybersecurity Risk Modeling for SCADA Systems," in *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4486-4495, Dec. 2018, doi: 10.1109/JIOT.2018.2822842.

[16] D. E. Kouicem, A. Bouabdallah, and H. Lakhlef, "Internet of things security: A top-down survey," *Computer Networks*, vol. 141, pp. 199–221, 2018.

[17] C. H. Liu, Q. Lin and S. Wen, "Blockchain-Enabled Data Collection and Sharing for Industrial IoT With Deep Reinforcement Learning," *in IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3516-3526, June 2019, doi: 10.1109/TII.2018.2890203.

[18] J. Huang, L. Kong, G. Chen, M. Wu, X. Liu and P. Zeng, "Towards Secure Industrial IoT: Blockchain System With Credit-Based Consensus Mechanism," *in IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3680-3689, June 2019.

[19] W. Liang, M. Tang, J. Long, X. Peng, J. Xu and K. Li, "A Secure FaBric Blockchain-Based Data Transmission Technique for Industrial Internet-of-Things," in *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3582-3592, June 2019.

[20] Q. Wen, Y. Gao, Z. Chen and D. Wu, "A Blockchain-based Data Sharing Scheme in The Supply Chain by IIoT," *2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS)*, Taipei, Taiwan, 2019, pp. 695-700.

[21] J. Wan, J. Li, M. Imran, D. Li and Fazal-e-Amin, "A Blockchain-Based Solution for Enhancing Security and Privacy in Smart Factory," in *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3652-3660, June 2019, doi: 10.1109/TII.2019.2894573.

[22] E. K. Wang, Z. Liang, C. M. Chen, S. Kumari, and M. K. Khan, "PoRX: A reputation incentive scheme for blockchain consensus of IIoT," *Future Generation Computer Systems*, vol. 102, pp. 140-151, 2020.

[23] Y. Lu, X. Huang, Y. Dai, S. Maharjan and Y. Zhang, "Blockchain and Federated Learning for Privacy-Preserved Data Sharing in Industrial IoT," in *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 4177-4186, June 2020.

[24] M. Shen et al., "Blockchain-Assisted Secure Device Authentication for Cross-Domain Industrial IoT," in *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 5, pp. 942-954, May 2020, doi: 10.1109/JSAC.2020.2980916.

[25] M. Y. Kubilay, M. S. Kiraz, and H. A. Mantar, "KORGAN: An Efficient PKI Architecture Based on Permissioned-Blockchain by Modifying PBFT Through Dynamic Threshold Signatures", *IACR Cryptol. ePrint Arch.*, 2019, 1141.

[26] W. Tsai, X. Bai and L. Yu, "Design Issues in Permissioned Blockchains for Trusted Computing," *2017 IEEE Symposium on Service-Oriented System Engineering (SOSE)*, San Francisco, CA, 2017, pp. 153-159, doi: 10.1109/SOSE.2017.32.

Design and Finite Element Analysis of IoT based Smart Helmet

Ninad V. Joshi
Student, School of Electronics and Communication
MIT World Peace University
Pune, India
joshininadv@gmail.com

Sumedh P. Joshi
Student, School of Mechanical Engineering
MIT World Peace University
Pune, India
sumedhj213@gmail.com

Malhar S. Jojare
Student, School of Electronics and Communication
MIT World Peace University
Pune, India
malharjojare0664@gmail.com

Neel S. Joshi
Student, School of Civil Engineering
MIT World Peace University
Pune, India
neeljoshi2699@gmail.com

Anjali R. Askhedkar
Assistant Professor, School of Electronics and Communication
MIT World Peace University, Pune, India
anjali.askhedkar@mitwpu.edu.in

Abstract*— As mentioned in the global two-wheeler market report, the two-wheeler market is projected to surpass $ 125 billion by 2024. According to WHO statistics, approximately 1.35 million people die each year as a result of road traffic crashes. And more than half of those are pedestrians, cyclists, and motorcyclists. Correct helmet use can lead to a 42 percent reduction in risk of fatal injuries and a 69 percent reduction in the risk of head injuries. These are concerning statistics. These growing numbers call forth new and upcoming technological advancements concerning the safety of the users. The main reason behind coming up with this smart helmet is the very issue of safety. The main objective of the project is to build a smart system which is integrated with sensors and modules for the safety of the user. While working on improving the safety of riders, it was ensured that its aesthetics and ergonomics are not compromised.

Keywords*— IoT, FEA, Smart Helmet

I. Introduction

Advancements in motorcycles have been faster than in helmets. Irrespective of whether one rides a road bike, mountain bike, or motorcycle, a good helmet is a must to stay safe. But helmets can turn uncomfortable after a sustained period of use. And navigating with the help of maps while having a helmet on is not very convenient. So basic concepts of IoT are used in the design.

A Bluetooth module enables connectivity to the smartphone. To ensure rider safety, vibration sensors detect collision and location information can be communicated. Other features also ensure rider safety. We had proposed a basic idea of our smart helmet which incorporated the user's safety as well as the user's needs.

For user safety, the impact sensor interfaced with the GPS module has been used to detect if the user comes across any accident. Texas Instruments kits are used for noise cancellation which is able to provide noiseless transmission between the users using Bluetooth module also providing the consumer a breath-taking surround-sound experience. The basic components of the helmet are Shell (Outer Layer of helmet), EPS Liner (High density foam to absorb shocks), Comfort Liner (Ensures comfort of head), Cheek Pads (for protection from side impacts), Visor (movable front shields to protect the face of the users), and a Smart Buckle system.

II. Literature Review

The goal of this literature survey is to compare the several approaches and methodologies that have been published in the field of smart helmet design.

Smart Helmet for Safe Driving: The use of FSR to detect whether helmet is worn by the user, MQ-3 alcohol sensor to detect whether the driver is drunk or not, vibration sensors to detect vibration and RF module for communication are the modules proposed in the work for smart helmet in [2].

A smart helmet using GSM and GPS with module SIM808, a limit switch to detect whether helmet is worn or not and accelerometer for continuous speed checking is proposed in a further work [3].

In reference [14], they have used FSR to check whether helmet is worn or not and accordingly the bike will start, a camera is placed in the bike which will check for the helmet. They have used MQ3 i.e. alcohol sensor to check the

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

permissible range of alcohol to start the bike and message will be sent to registered number. GPS system in bike will track the location and send message if injured

The research paper [4] they used a limit switch for checking of helmet worn by the user or not, accelerometer sensor for keeping track of speed and if exceeding the ignition gets cut, GSM module to communicate the message. [5] provides the basic insights regarding the oblique impact of the motorcycle helmet and the angular acceleration of the head foam due to various factors.

[6] Provides vital information about the conventional design of the helmet and the benefits of using helmets along with a statistical case study of accidents in Thailand.

Research paper [11] gives the basic understanding of the safety limits for impact testing for helmets as per the BIS (Bureau of Indian Standards). In the study conducted in [13], a car accident was reconstructed in order to simulate and analyze the motion, impact velocity, impact angle in the collision.

Based on the current evidence, using smart helmets helps in reducing injuries as numerous studies are introduced by different authors all over the world and these innovations will prove beneficial for the people.

III. Market Analysis in India

According to a sample survey conducted by us in Pune, India, 25% of the people preferred cabs for daily commuting. This was linked to the Ola, Uber boom in India. 15% people used 4-wheelers for commute. 10% prefer public transport and the rest use mopeds and bikes. About 50% of people that we surveyed used 2- wheelers, which was due to the fact that 2-wheelers are cheaper, compact, and easily maneuverable. Due to the decisions of the Government of India to improve and focus on road safety, it is mandatory for 2-wheeler users to wear helmets, failing which can attract hefty fines. In spite of the introduction of these strict traffic regulations, 5% of drivers did not wear helmets. Many also felt the need for an inbuilt navigation system because using smartphones for navigation can sometimes lead to accidents and isn't that reliable. They also complained regarding the bulky structures of the helmets and acute neck and back pain. The need for an ergonomic helmet was felt. Another big problem is about the post-accident situation. Very few people readily come forward to help an injured biker in India which is the harsh reality. According to another national survey, about 37 million 2-wheelers run on the roads in India. Hence, the necessity of incorporating an accident detection and communication system in helmets is crucial.

IV. Working of the Circuit

The block diagram explains the basic functions of the various modules of the Smart Helmet system. As safety is of paramount importance in our product, it is designed in such a way that the helmet's features would not start up until the rider is buckled in. The capacitive sensor is placed inside the buckle of the helmet. So only when the user ties the strap and locks the buckle of the helmet will the sensor get actuated. This idea ensures that the users take the basic safety.

The SW-420 sensor will be continuously sensing and if the user meets with an accident, that is vibrations frequency is greater than the preset threshold frequency (40 Hz), then the controller board will become high and the signal will be sent to the GSM module. Fig. 01 shows the block diagram of the system.

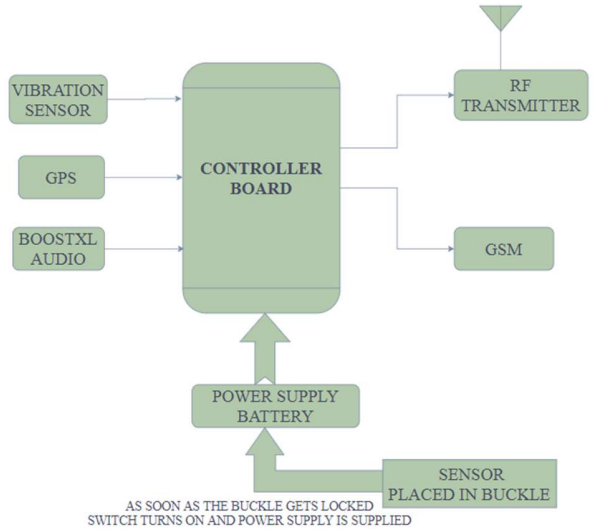


Fig. 01 - Smart Buckle System Block Diagram

A message of the user's location with the help of GPS module will be sent to the desired port via use of the GSM module. The port will receive this message in the form of an NMEA string. The flowchart is shown in the Fig. 02.

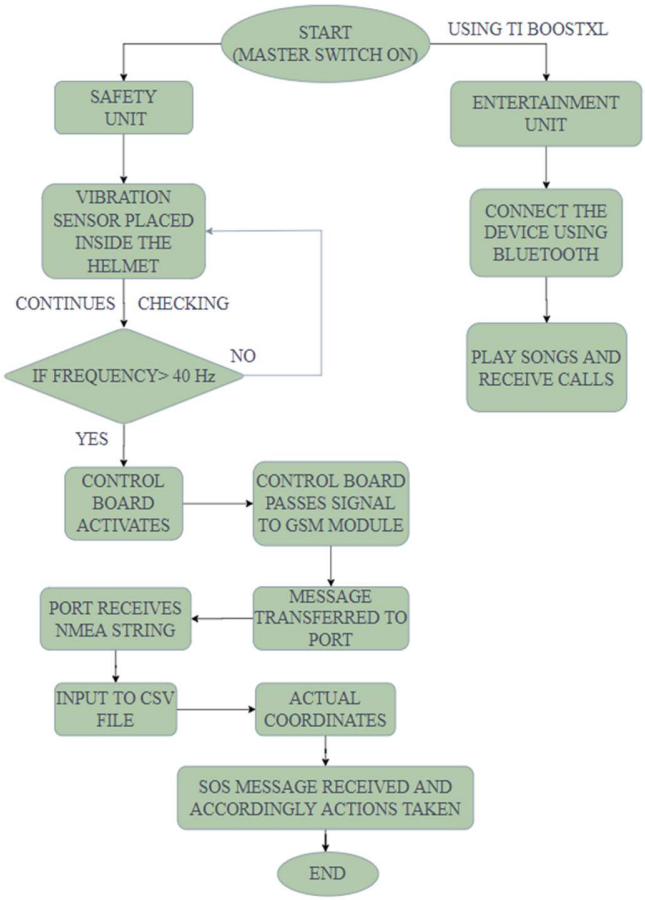


Fig. 02 – Flowchart of the System

V. Implementation Setup

A. Components Required

1) SW-420 Vibration Sensor
2) NEO-6MV2 GPS Module
3) Controller Board
4) SGW 8130 BT sensor tag
5) GSM Module (SIM900A GSM Modem)
6) TI BoostXL

B. SW-420 Vibration Sensor

The vibration sensor is shown in the Fig. 03

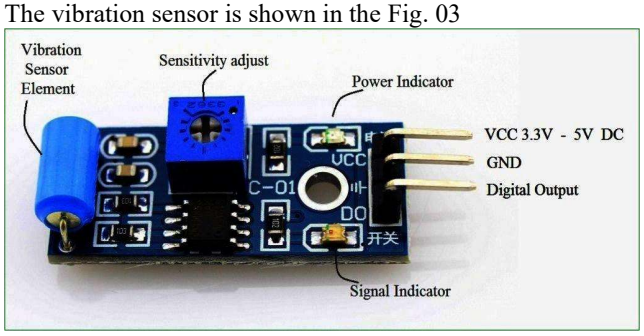


Fig. 03 – SW-420 Vibration Sensor

Vibration Sensor SW-420 is a highly sensitive non-directional sensor used to sense the vibrations which are placed inside the helmet. The position of vibration sensors is selected to be at the back side and are placed in such a way that the accuracy as well as the efficiency would be maximum. Placement is done such that when placed it would form an equilateral triangle with side length at 50mm all over the helmet. The sensor is interfaced with the controller board and GSM module and the threshold frequency of vibrations set to 40 Hz. The sensor would be continuously sensing the vibrations, it would be working in IF ELSE condition i.e. if the sensor senses frequency greater than the desired or set frequency a message would be sent via GSM to the nearest port and necessary actions would be taken. And in other conditions it would recheck and sense continuously. The size of this module is 30mm x 15mm.

C. NEO-6VM2 GPS Module

Fig. 04 shows the actual GPS module.

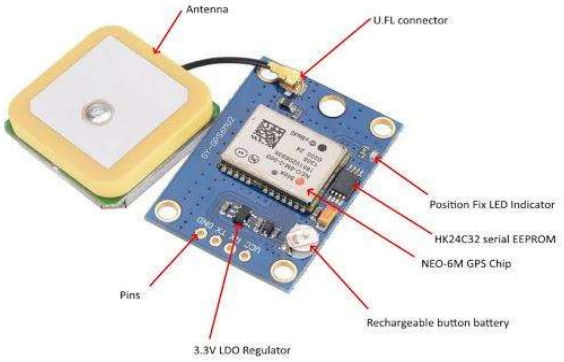


Fig. 04 – NEO-6MV2 GPS Module

The NEO-6MV2 is a GPS (Global Positioning System) module and is used for navigation. The module simply checks its location on earth and provides output data which is longitude and latitude of its position. The NEO-6VM2 GPS Module is interfaced with a controller board occupying two pins. Vibration sensors would perform their task and if a message is to be delivered then only this module would come into action. The interfacing between the GPS module and the controller board is performed by using a +5V supply from the power side to the controller board and any ground pin. Any two pins would work for the serial communication. The GPS receiver module gives output in standard NMEA i.e. National Marine Electronics Association string format. It provides output serially on a Tx pin with default 9600 baud rate. This NMEA string output from a GPS receiver contains different parameters separated by commas like longitude, latitude, altitude, time etc.

D. SGW 8130 BLE sensor tag

The SGW 8130 sensor tag is a highly integrated Bluetooth Low Energy (BLE) device with different sensor options. It reports temperature, humidity, light intensity and motion data through a wireless Bluetooth 4.0 link. Bluetooth sensor tag is interfaced with the controller board. The Booster audio module provided by TI is also interfaced with the Bluetooth module. It would be an additional feature included in our product. Music will be played when required by the user and other times it will guide the user with directions to the required location. Another feature of this product is for biker gangs which would help them for a faster and easier way of communication.

E. SIM900A GSM Modem

Fig. 05 represents the GSM Module.

Fig. 05 – SIM900A GSM Module

The SIM900A is a complete Dual-band GSM/GPRS solution in a SMT module which can be embedded in the customer applications. Featuring an industry-standard interface, the SIM900A delivers GSM/GPRS 900/1800MHz performance for voice, SMS, Data, and Fax in a small form factor and with low power consumption. GSM digitizes and compresses data, then sends it down a channel with two other streams of user data, each in its own time slot. GSM is used for transferring the SOS message.

F. TI BoostXL

Texas Instruments BoostXL is a high-quality audio playback with onboard 14-bit DAC (can be bypassed if target MCU has an integrated DAC). It can automatically switch from onboard speakers and microphone to a headset with microphone where

onboard microphone supports sampling rates of up to 20kHz.This module is used as a part of our Entertainment Unit. Fig. 06 shows the detailed circuit board of the BoostXL.

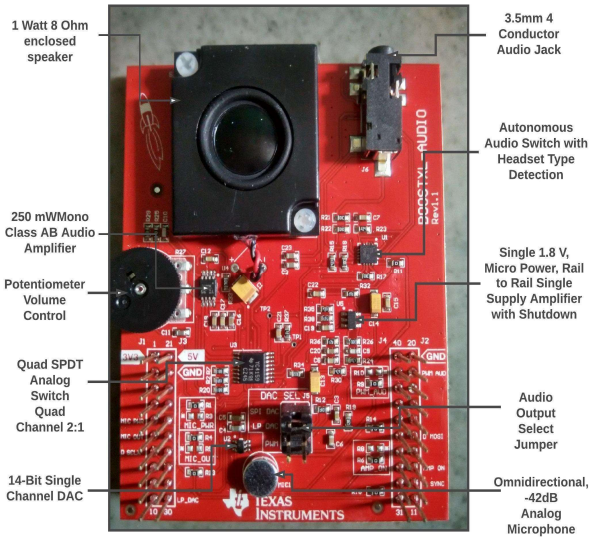


Fig. 06 – TI BoostXL

When plugged into a LaunchPad Development Kit, the BOOSTXL-AUDIO Audio Booster Pack Plug-in Module adds audio input functionality from a microphone, as well as audio output through an on-board speaker. Headphone input and output is also supported, and automatically enabled when a plug is inserted into the Booster Pack. This audio I/O stream lets the user experiment with the digital signal processing (DSP) and filtering capabilities of the microcontroller found on the attached LaunchPad Development Kit.

VI. Circuit Diagram

Fig. 07 shows the circuit diagram of the safety unit.

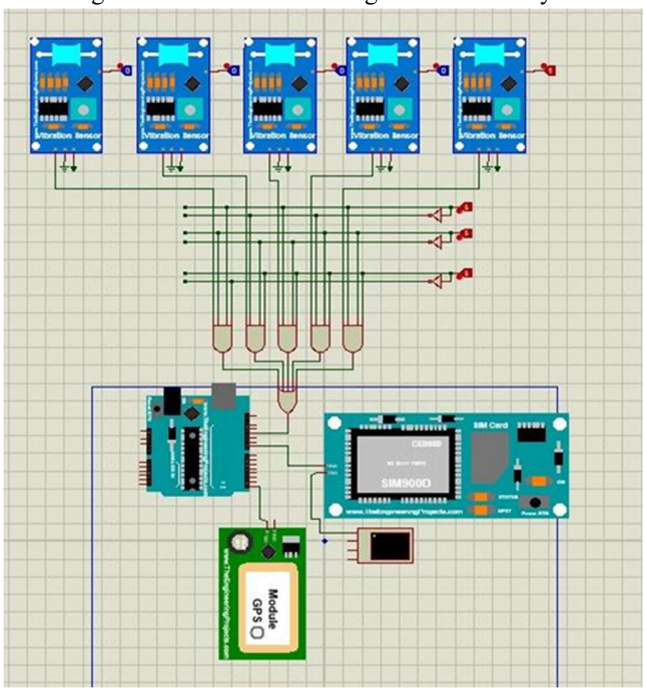


Fig. 07 – Circuit Diagram of the Safety Unit

VII. CONSTRUCTION OF THE HELMET

This capacitive sensor would be the most important core component of the system as it would act as an input to the power supply, because of which the entire system would start. So, if the rider simply puts on the helmet without locking, he/she will not be able to access the various features other than the basic impact protection. This idea ensures that the users take the basic safety precautions prior to driving, thus instilling in their minds the necessity of safety on roads.

After the buckle is locked, the Safety Unit and Entertainment Unit activate. The smartphone can be paired with the SGW 8130 BLE sensor tag and the rider can enjoy music and receive calls. The TI BoostXL module provides excellent sound quality for this purpose. In the Safety Unit, the SW-420 Vibration sensors continuously sense the external vibrations. They are preset at 40Hz of threshold vibration frequency. If the rider meets with an accident, vibrations of frequency more than 40Hz are produced during the impact and the sensors activate the control board. The control board then passes a signal to the GSM module and the GPS module. The GPS module detects the location and sends an NMEA string to the control board. This NMEA string serves as an input to the CSV file conversion algorithm written in Python language and the global coordinates of the current location are obtained. This data is then broadcasted in the form of an emergency SMS through the GSM module to the port designed so that necessary actions are taken.

VIII. Module Placement

The sensor placement is done according to the space available between the outer shell and the inner foam layer. The placement is designed based on the individual sensor sizes. The Fig. 08 shows the module placement briefly inside the helmet.

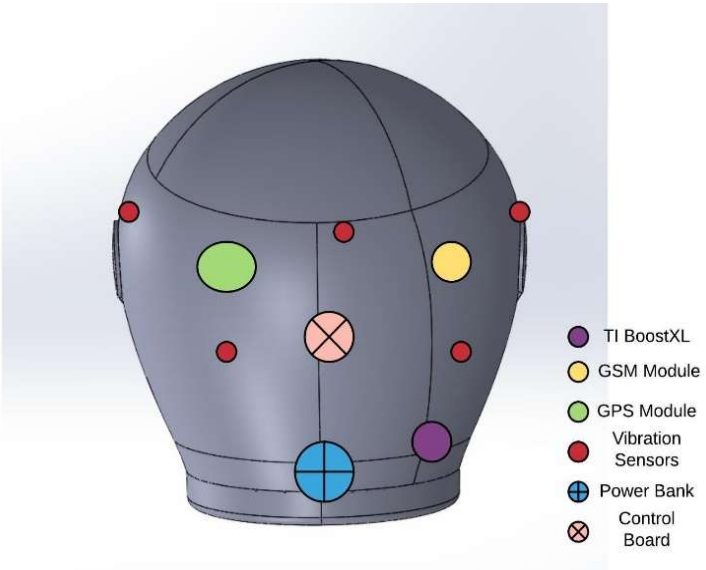


Fig. 08 - Sensor and Module Placement (Rear View)

Here the sensors and modules are placed in the rear part of the helmet. These modules are embedded in a foam casing to safeguard them against vibrations and other foreign entities. Fig. 09 represents the flowchart of the code.

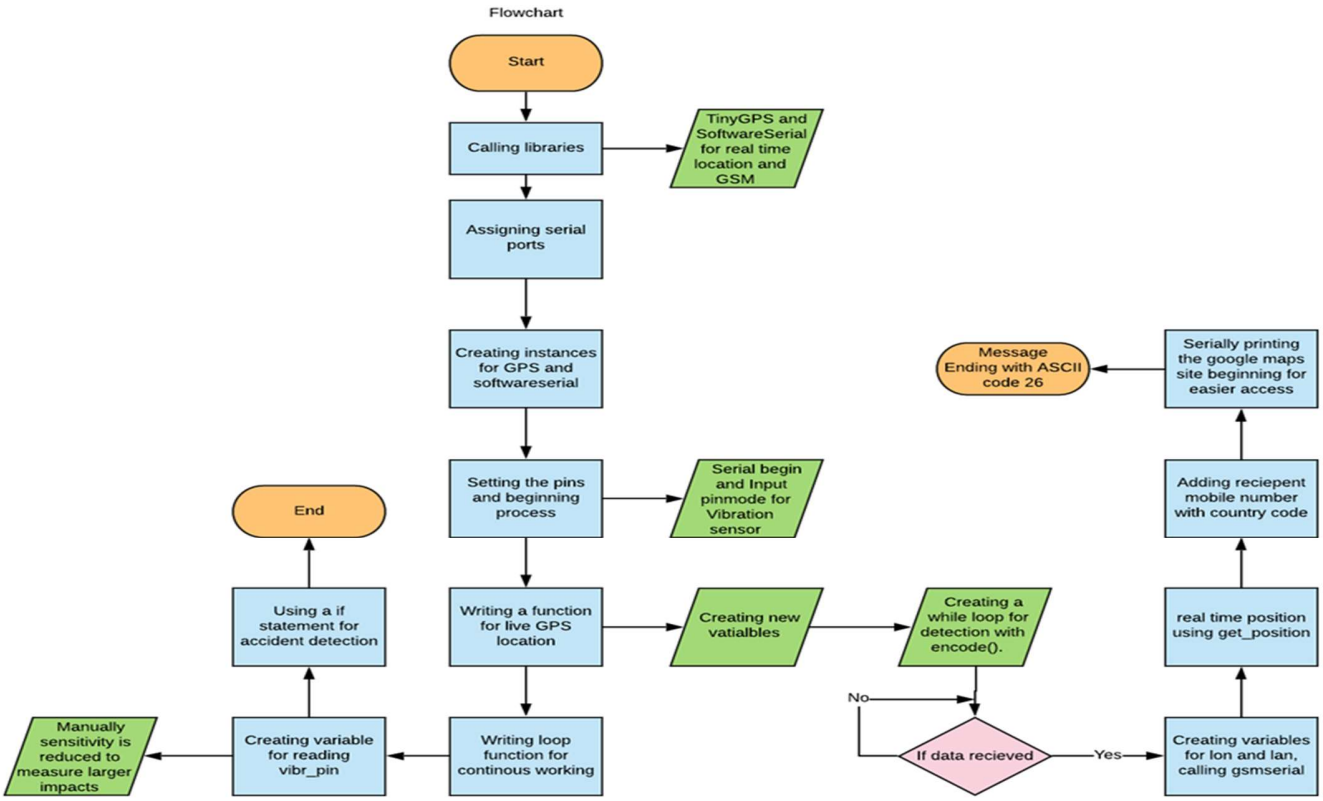


Fig 09. Flowchart of code

IX. Replaceable Battery System

The source of power for the helmet will be replaceable battery packs. 2 battery packs will be provided. The users will be able to recharge one pack after its power dies down and can replace it with the second one. This will provide uninterrupted usage of the helmet, in that the user will not have to wait till the battery recharges and can quickly switch over to the other pack.

X. Advantages

The detection feature is very useful due to which immediate and essential, prompt, and speedy help can be given. The Bluetooth is enabled for music as well as receiving and making calls. The Smart safety buckle ensures that the rider is wearing the buckle at all times as it is essential to hold the helmet in place during an accident. This system does not require any smart-phone app and can be accessed from any device with Bluetooth capability.

The helmet is cheaper than other smart helmets in the market and offers various extra features and additionally, it is considerably light and weighs only 1723 gm including all the electronic components. The smart safety buckle system eliminates the false activation of the system when the helmet accidentally falls from the hands while carrying. The replaceable battery system is easy to use and is practical.

XI. Simulation

The safety system was simulated by applying real-time accident constraints and conditions.

The circuit is simulated on Proteus Software, and the SOS message obtained with the location of the accident in the form of a Google Maps link. The circuit simulation is shown in Fig. 10.

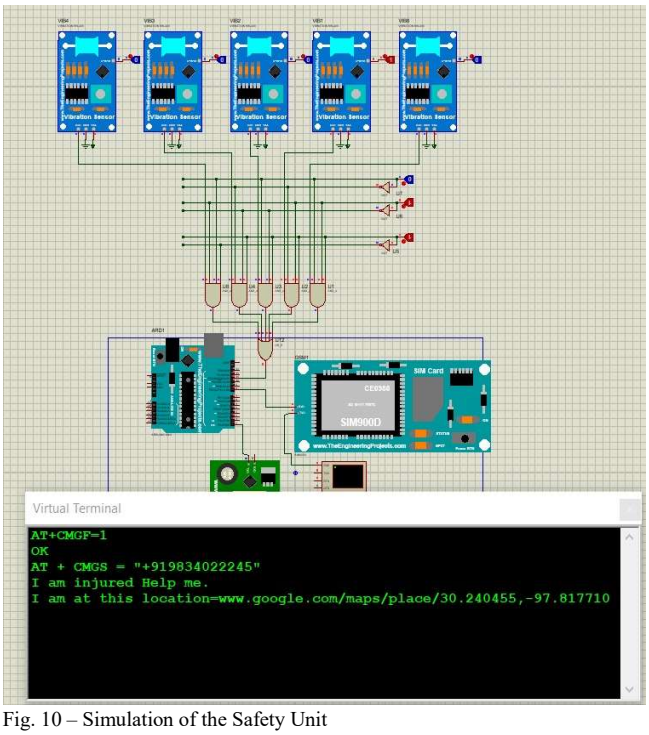


Fig. 10 – Simulation of the Safety Unit

XII. Design and FEA of Helmet

A. Components Required

The helmet body was designed using SolidWorks 2018 software and the materials were selected based on their strength, weight and durability. Dimensions of the helmet are – 217mm×193mm×202mm (length, breadth and height respectively). Fig. 11 shows the CAD design of the helmet.

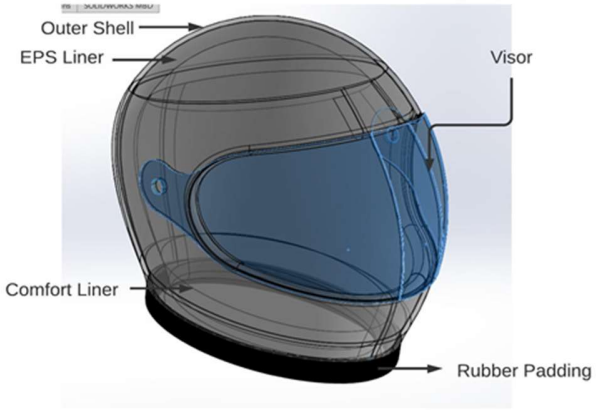


Fig. 11 – CAD design of the helmet

The Fig. 12 represents the helmet part and the materials used. Different layers and parts were designed including the Outer Shell which forms the outermost protective layer of the helmet, EPS Liner or the layer under the outer shell, a comfort liner which makes up the innermost layer that comes in contact with the rider and a rubber padding for perfect fitting.

With all the sensors and modules placed, the ready to use helmet weighs around 1723 gm which is light compared to other normal helmets where the average weight range is around 1600-1800 gm.

TABLE I
MATERIAL PROPERTIES OF HELMET COMPONENTS

Helmet Part	Material	Properties
Outer Shell	Polycarbonate	Young's Modulus = 2.1 GPa Density = 1.22 g/cm^3 High Strength to Weight Ratio
EPS Liner	Extended Polystyrene	Young's Modulus = 3 GPa Density = 0.64 g/cm^3 Ductile, Excellent Damping Properties
Comfort Liner	Polyurethane Foam	17% Spandex, 83% Nylon (40 denier) Breathable, Soft, Lightweight
Cheek Pads	Polyurethane Foam	Aerospacer Mesh Enhances Airflow
Visor	Acrylic	Durable and Excellent Light Distortion Property
Buckle Straps	Nylon	Excellent Strength and Durability
Rubber Padding	Vulcanized Rubber	Excellent Stiffness and Durability

Fig. 12 – Material Properties of Helmet Components

B. FEA of the Helmet

Finite element analysis was performed and the corresponding forces and constraints were applied on the system to simulate an impact situation similar to a real-life accident. A 19.5kN load was applied to the outer shell of the helmet as per the BIS (Bureau of Indian Standards) standards for helmets and the results were recorded. A tetrahedral mesh was generated and the analysis was performed. The Fig. 13 shows the mesh generation.

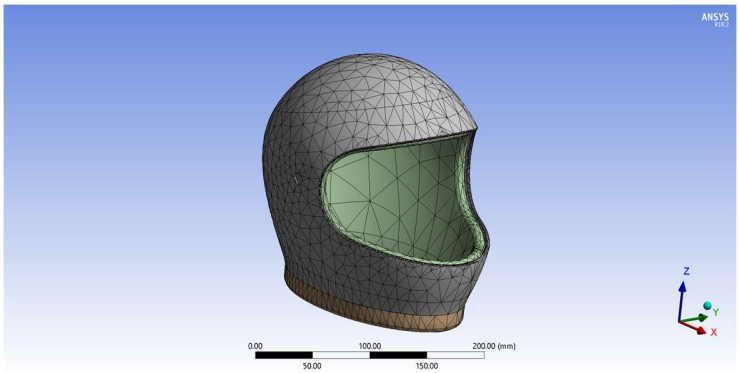


Fig. 13 – Mesh Generation for FEA

The static structural deformation analysis was performed using the ANSYS Workbench 2018.2 software. Fig. 14 shows the static deformation analysis for the top portion loading.

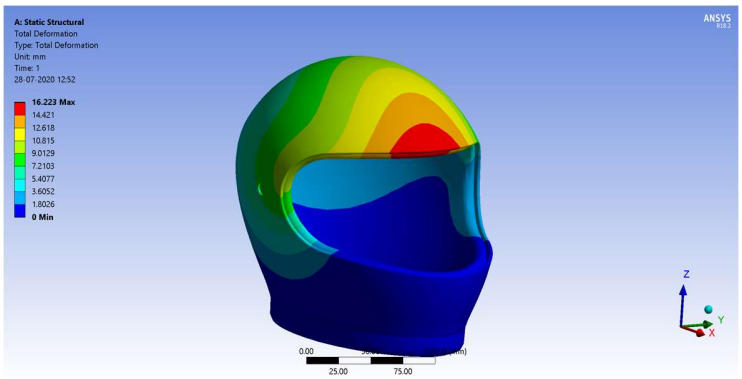


Fig. 14 – Static Deformation Analysis (19.5 kN applied to the top portion)

The maximum deformation was found to be 16.2 mm which is under the acceptable limit of the BIS standard which lies between 6 - 24mm This result showed the robust and strong design of the outer shell of the helmet.

In the second case, 19.5 kN load was applied to the side portion of the helmet body and the results were recorded. The maximum deformation was found to be 9.032 mm which lies in the acceptable limit of the BIS standards. The Fig. 15 shows the static deformation analysis for the side portion loading.

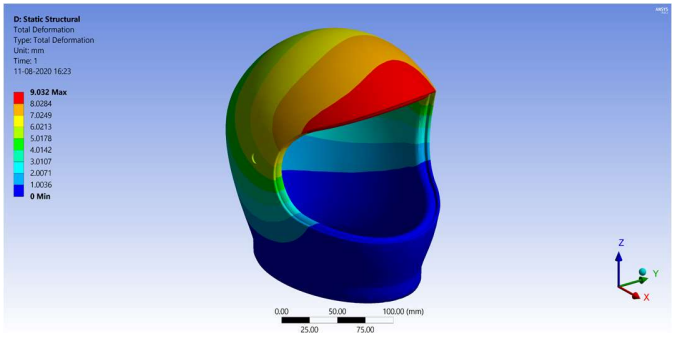


Fig. 15 – Static Deformation Analysis (19.5 kN applied to the side portion of the helmet)

For this iteration, the Equivalent Stress and Strain percentage were found to be 157.28 MPa and 4.766% respectively. Fig 16 shows the equivalent stress analysis. Whereas, the strain analysis is shown in the Fig. 17.

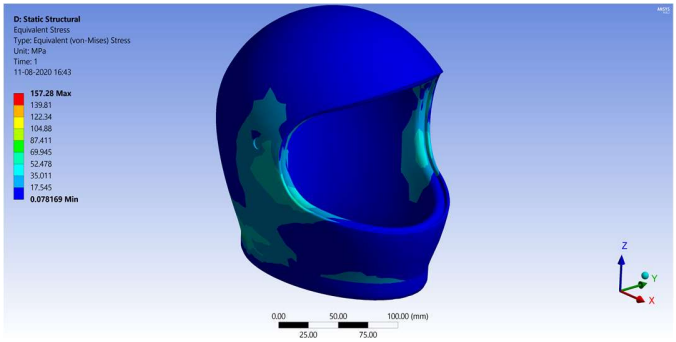


Fig. 16 – Equivalent Stress Analysis

Comparing these results to the study published in [11], the stress and strain values are under the acceptable limit.

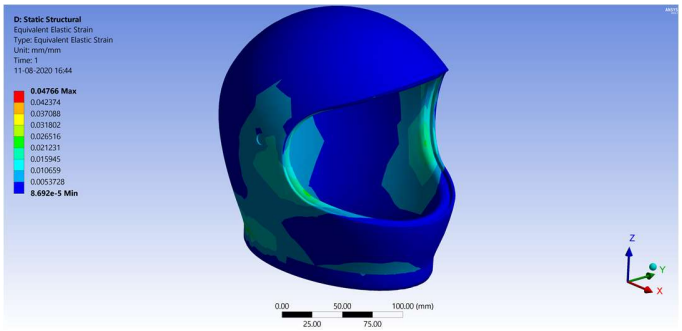


Fig. 17 – Equivalent Strain Analysis

Dynamic analysis of the impact was also performed where the helmet was given a velocity of 10 m/s and the side impact condition was simulated. A concrete block was used for this impact. This analysis was used to predict the real-life situation of an accident and the extent of damage the helmet body would suffer. The impact analysis condition is shown in Fig. 18. In this figure, the yellow body is the entire helmet and the concrete block can be seen resting as a rigid body. The material properties were given to the parts and the analysis was carried out.

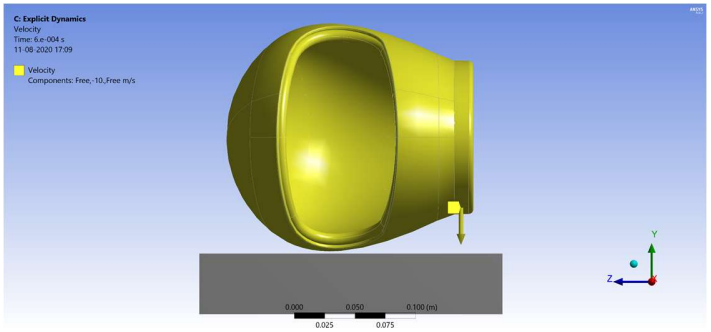


Fig.18 – Impact Analysis Condition

The stress values were calculated and the maximum stress was found to be 2.75e9 Pa and can be seen in Fig. 19.

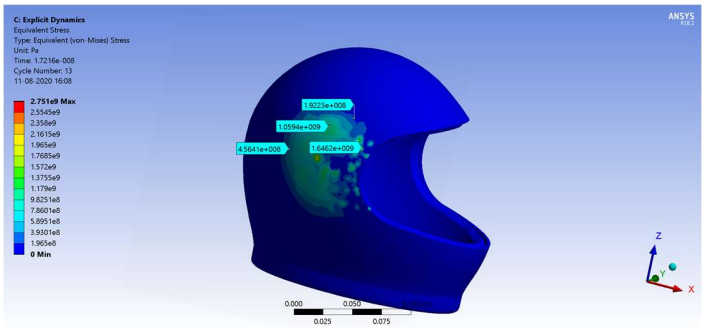


Fig.19 – Dynamic Analysis (Equivalent Stress)

From the above results, it was found that the design of the helmet was safe and the deformation values were under the acceptable limit of the BIS standard for helmet testing.

XIII. Product Comparison

The following comparison table given in Fig.20, compares our idea with five other products already in existence. They have been compared as follows -

TABLE II
PRODUCT COMPARISON WITH EXISTING PRODUCTS

Product	Bluetooth	GPS	Accident Detection	Cost	Weight (gm)
Our Product	YES	YES	YES	$225	1723
Headfox MX	YES	YES	NO	$115	1830
Greenstone G6	YES	YES	NO	$135	2320
Quinn Spitfire Rosso	YES	YES	YES	$400	1750
ILM Bluetooth Integrated Helmet	YES	YES	NO	$192	2000
Harley Davidson Boom! N02	YES	YES	NO	$550	1680

Fig. 20 – Product Comparison with Existing Products

XIV. Future Scope

This system is designed keeping in mind the road safety norms and convenience. The current system is designed to accept single emergency contact but in the future a provision can be made to store multiple emergency contacts. Further, this system can be programed to send the SOS response to the nearby medical service by incorporating AI and Google cloud services to enhance the emergency response mechanism. With the advent of 5G, a speaking assistant can be embedded into the system itself. Further, an in-built solar powered battery system where the solar panels will be placed on the portion of that helmet that experiences least damage such as the rear portion can also be incorporated to improve the functionality.

References

[1] Techsci Research, "Global two-wheeler market by vehicle type (scooter/moped and motorcycle) by engine capacity (up to 125cc, 126-250cc, 250-500cc and above 500cc), by region, competition, forecast & opportunities," Research and Markets, ID. 4769671, April 2019.

[2] K. Shravya, Y. Mandapati, D. Keerthi, K. Harika and R. Senapati, "Smart helmet for safe driving," ES3 web of Conferences, vol. 87, January 2019.

[3] A. Varade, N. Gajbhiye, Mousam, V. Panchbhai, "Smart Helmet using GSM and GPS," International Research Journal of Engineering and Technology, vol. 04, pp. 1662-1667, March 2017.

[4] T. Hadke, M. Kharwade, P. Nagarkar, V.Panchbhai, "Intelligent Helmet Band," International Research Journal of Engineering and Technology, vol. 04, pp. 2526-2530, June 2017.

[5] N. Mills, S. Wilkes, S. Derler, A. Flisch, "FEA of oblique impact tests on a motorcycle helmet," International Journal of Impact Engineering, vol. 36, pp. 913-925, July 2009.

[6] T. Teng, C. Liang, V. Nguyen, "Development and validation of finite element model of helmet impact test," Proceedings of the Institution of Mechanical Engineers Part L Journal of Materials Design and Applications, vol. 277, pp. 82-88, January 2013

[7] J. Wobrock and T. Smith, V. Kasantikul, W. Whiting, "Effectiveness of collision-involved motorcycle helmets in Thailand," Annual Proceedings Association for the Advancement of Automotive Medicine, vol. 47, pp. 1-23, 2003.

[8] N. Manjesh, S. Raju, "Safety measures for two wheelers by smart helmet and four wheelers by vehicular communication," National Conference on Developments, Advances & Trends in Engineering Sciences, 2015, pp. 15-19.

[9] P. Chitte, S. Akshay, A. Thorat, N. Bhosale, "Smart helmet & intelligent bike system," International Research Journal of Engineering and Technology, vol. 03, pp. 483-487, May-2016.

[10] P. Li, R. Meziane, M. J. -. Otis, H. Ezzaidi and P. Cardou, "A smart safety helmet using IMU and EEG sensors for worker fatigue detection," 2014 IEEE International Symposium on Robotic and Sensors Environments (ROSE) Proceedings, Timisoara, 2014, pp. 55-60.

[11] V. Gandhi, R. Kumaravelan, S. Ramesh, M. Venkatesan, "Performance analysis of motor cycle helmet under static and dynamic loading," Mechanics and Mechanical Engineering, vol. 18, pp. 85-96, May 2014

[12] H. Kingsbury, P. Rohr, "Structure characteristics of motorcycle helmets," SAE International Congress and Exposition, 1981, pp. 1-20.

[13] J. Chen, J. Yang, Z. Xuenong, Z. Chao, Y. Xiaoping, L. Weiqiang, W. Meichuan, "FEA of helmet-head injury protection based on motorcycle accident reconstruction," 2013 Fifth International Conference on Measuring Technology and Mechatronics Automation, Hong Kong, 2013, pp. 570-573.

[14] K. Vidhya, M. Kasiselvanathan, "Smart Helmet and Bike System," International Journal of Recent Technology and Engineering, vol. 07, pp. 1-4, December 2018

Jig Detection Using Scanning Method Base On Internet Of Things For Smart Learning Factory

1st Ridho Hendra Yoga Perdana
Department of Electrical Engineering
State Polytechnic of Malang
Malang, Indonesia
ridho.hendra@polinema.ac.id

2nd Nurul Hidayati
Department of Electrical Engineering
State Polytechnic of Malang
Malang, Indonesia
nurulhid8@polinema.ac.id

3rd Ahmad Wilda Yulianto
Department of Electrical Engineering
State Polytechnic of Malang
Malang, Indonesia
ahmadwildan@polinema.ac.id

4th Vipkas Al Hadid Firdaus
Department of Information Technology
State Polytechnic of Malang
Malang, Indonesia
vipkas@polinema.ac.id

5th Nila Novita Sari
Department of Electrical Engineering
Politeknik Negeri Bandung
Bandung, Indonesia
nila.novita@polban.ac.id

6th Dodit Suprianto
Department of Information Technology
State Polytechnic of Malang
Malang, Indonesia
dodit.suprianto@polinema.ac.id

***Abstract*—Nowadays, the evolution of a factory has pushed towards a smart learning factory. By having the ability to learn, monitor each device and the resulting product will get more precise product results according to design. Detection of a place to put a product or called a jig on a shuttle whose number is different becomes something that is often missed from observation. With the scanning method with five pieces of ultrasonic sensors, the detection speed is 50 µs, and the detection accuracy is 100% for distances less than 25cm. The detection process of the jig is sent and stored by IoT Gateway as a Big Data Cluster via wifi media with a performance of 99.4%. The process of storing data on the IoT Cloud as the Main Big Data has a performance of 100% of the data on the IoT Gateway.**

Keywords—Smart Learning Factory, Internet of Things, Jig, Scanning, Ultrasonic

I. INTRODUCTION

The development of smart city now has pushed the implementation of an industrial revolution called the industrial revolution 4.0[1], [2]. Therefore in the industrial world is also experiencing very rapid development, with the implementation of the industrial revolution 4.0, the industry is competing to change the automation which was previously only in the form of electronic automation to be internet-based[1]. The concept of the internet offered in Industry 4.0 is the concept of the Internet of Things (IoT). By using the Internet of Things all equipment and systems contained in the industry will be monitored in real-time, it is hoped that the application of the Internet of Things technology can increase production in the industry[2]–[5].

Smart learning factory is a system in the industry that can adapt to changes in production and is studied to produce results that are following standards [6], [7]. In the smart learning factory, there are several layers, and the first is the mobility layer, location base service and cloud-based service[8]. Mobility layer is a layer that functions for control and sensing in the course of employment in the industry. The location-based service layer functions as a network provider so that the nodes in the mobility layer can be connected to the cloud layer. Whereas the cloud-based service layer functions to store all data read from the node[9]. The use of a microcontroller with wireless communication is used to support the needs of the mobility layer[10]. While at the location-based service layer, many use the WIFI network to support faster communication[6]. At the cloud layer, a large deviation is needed because of the large amount of data sent by each node in an industry, so big data is used to overcome it. The use of big data at this time is very important for data storage, especially in the smart learning factory, because by storing data history, the system can learn events that exist in the previous production[7], [9]. However, the use of big data in the cloud has a weakness that is if the internet connection is disturbed then the data will not arrive. The use of raspberry pi, which functions as a big data cluster is an alternative solution in the event of an interruption in internet connection[11]. However, if only using big data in clusters, the industry cannot be monitored from other regions.

Smart Learning Factory is a system that can connect all production equipment in a factory and even more than one factory. Everything is connected and integrated to achieve fast production results and with the best results[12]. In a Smart Learning Factory, there is a problem in detecting jigs or where to place a product in the production process. The use of ultrasonic sensors with scanning methods can improve accuracy in terms of readings while also reducing the interference effect between ultrasonic sensors[13]. Nevertheless, ultrasonic sensors have never been used to detect and count a jig.

In this research, a device that can detect and calculate jigs is proposed. By implementing the scanning method and ultrasonic sensor on a node, it can make readings better. Internet of Things is used for sending data to the cloud with a gateway using raspberry pi as its big data cluster so that it is expected to improve product yields that are getting closer to the standard.

978-1-7281-5335-3/19/$31.00 ©2019 IEEE

II. METHOD

A. Smart Learning Factory

Smart Learning Factory is a change from factories whose supporting equipment can be connected to the internet called the industrial revolution 4.0[12]. Figure 1 shows an industry 4.0. In addition to being connected to the internet, the system can also learn the history of a product that has been produced. By studying the results of a product that has been made, it will produce a product that is more precise and in accordance with the desired design. In addition, the purpose of the smart learning factory is also to streamline work so as to increase the profit of a factory.

Fig. 1. Industry 4.0

B. Internet of Things

Internet of things technology can connect various things and equipment in an internet network. IoT technology has now been adopted in various elements[2], [4], [5], [7]. From a report published by business insider mentioned IoT devices would increase rapidly to 41 billion in 2027[14]. So that the industrial world is also experiencing a change towards the internet of things, the industrial world calls it the Industrial Internet of Things (IIoT)[5].

C. Jig Loading

Jig loading becomes an important part of the smart learning factory. The process that is done in the loading jig is to put the jig on a shuttle. Figure 1 shows a loading jig in a smart learning factory. Jig loading process must be by the position and timeliness on the shuttle. Otherwise, there will be a position mismatch, and an error occurs.

Fig. 2. Jig Loading

D. System Design

The system design proposed in this study is shown in Figure 3.

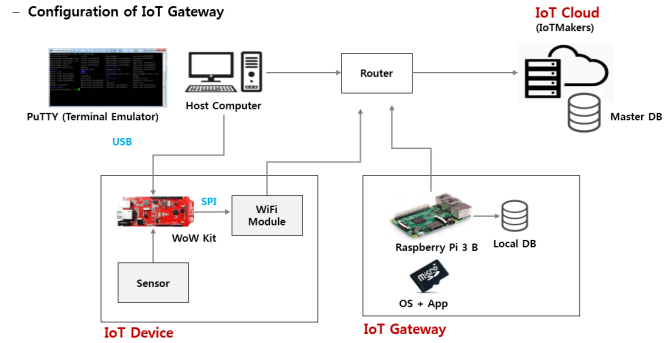


Fig. 3. system design

In picture 3, there are three main parts, namely the IoT Device, IoT Gateway and IoT Cloud. The first part is the IoT Device, and there are three devices, a microcontroller that functions as a device that processes data. An ultrasonic sensor that serves to detect jigs, and the last is a wifi module that functions as a communication module. Part of the IoT Gateway is a Raspberry Pi device that functions as a Big Data cluster. The two parts are connected through a wifi router that is also connected to the internet. The process of configuring the IoT Device and IoT Gateway uses a Host Computer that is also connected to the Wifi Router.

The jig detection process is carried out using an ultrasonic sensor on an IoT device. Then the data is sent by the microcontroller using a wifi network to the IoT Gateway with a raspberry pi device. In the IoT Gateway, which functions as a big data cluster, data will be stored on the local side. Once saved, the raspberry pi will send data to the IoT Cloud as the main storage.

Schematic diagram that is used in the scanning process is shown in Figure 4.

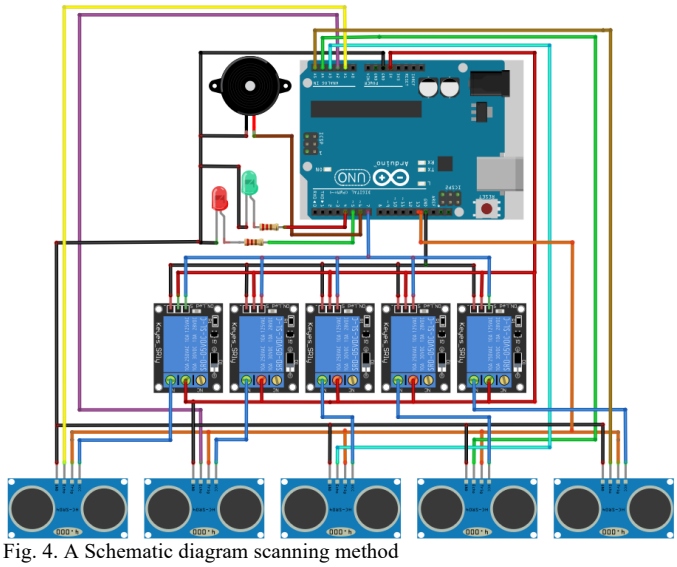


Fig. 4. A Schematic diagram scanning method

Figure 4 shows the ultrasonic sensor scanning method used. There are five ultrasonic sensors connected in parallel. The placement of the adjacent ultrasonic sensor, as shown in Figure 5 and Figure 6 shows the loading jig position, requiring the use of scanning methods to avoid interference between sensors.

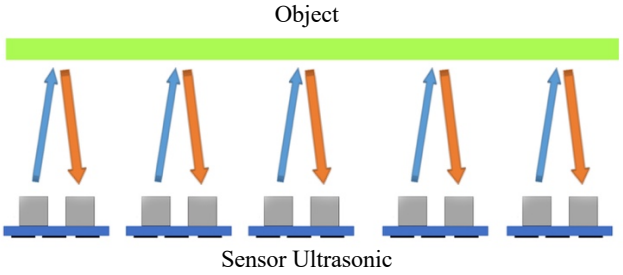


Fig. 5. Placement ultrasonic sensor

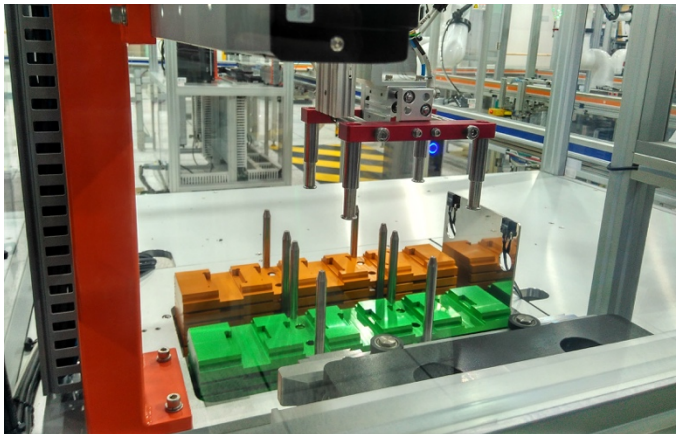


Fig. 6. Jig loading position

The number of sensors used to perform the scanning method is five units with relay control. Each relay is active alternately in the order shown in Figure 7.

time 1	**relay 1**	relay 2	relay 3	relay 4	relay 5
time 2	relay 1	**relay 2**	relay 3	relay 4	relay 5
time 3	relay 1	relay 2	**relay 3**	relay 4	relay 5
time 4	relay 1	relay 2	relay 3	**relay 4**	relay 5
time 5	relay 1	relay 2	relay 3	relay 4	**relay 5**

Fig. 7. sequential relay activation

The relay activation sequence shown in Figure 5 starts with time 1st to time 5st. With relay 1 being active first and then relay 2,3,4, and 5 is inactive. If relay 5 is active moreover relay 1,2,3,4 is inactive. The sequence repeats continuously until the device is turned off. Relays are not allowed to be active simultaneously. This will result in interference between sensors.

III. Result and Discussion

Tests conducted in this study were conducted with two events, namely partial testing and system testing. The testing begins with testing of ultrasonic sensors. The results obtained in the ultrasonic sensor test are shown in Figure 8.

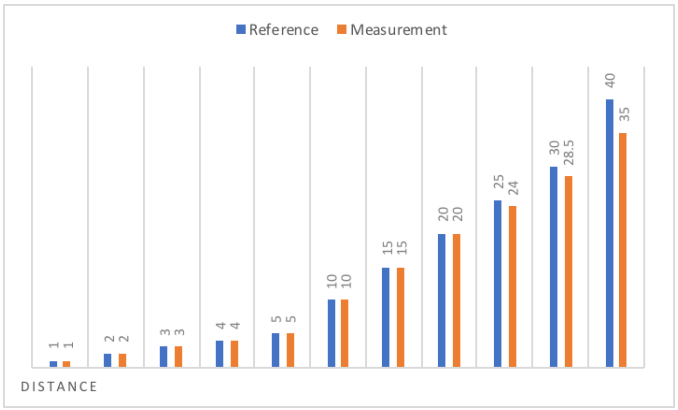


Fig. 8. Sensor ultrasonic testing result

In the distance test, the results obtained, as shown in Figure 8. Shown the ultrasonic sensor has very good accuracy at a distance of less than 25 cm. When the distance is more than 25 cm, the accuracy of the ultrasonic sensor has decreased.

The next test is testing the sensor reading time by scanning method. The results of testing the scanning method are shown in table 1.

TABLE I. Testing The Scanning Method

Time (µs)	Result
1	not detected
2	not detected
3	not detected
4	not detected
5	not detected
6	not detected

Time (µs)	Result
7	not detected
8	not detected
9	not detected
10	detected
11	detected

Table 1 shows the test starts from a time delay of 1 µs to 11 µs. Testing with a delay of fewer than 10 µs, the sensor cannot detect the object being detected. This is because the scanning process is too fast. So that with the minimum delay obtained by 10 µs, then in one cycle has a delay of 50 µs.

The implementation of the jig detection device that has been made is shown in Figure 9.

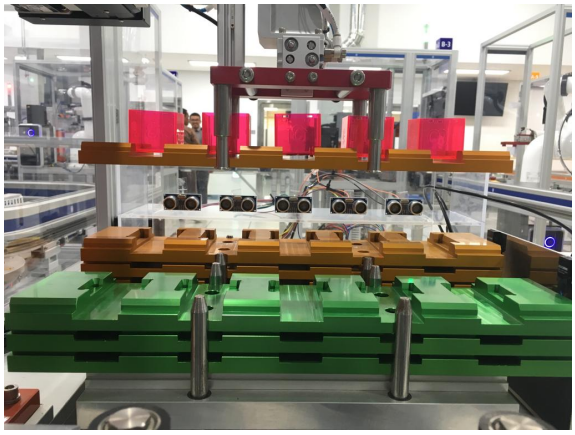


Fig. 9. Jig Detector Implementation

The jig detector is placed on the vertical side of the jig stack. That is because there is a robot arm that is on the top side of the jig. The robot arm serves to lift the jig from storage and put it on the shuttle.

The process of sending data to the IoT Gateway using a local network obtained the test results shown in Figure 10.

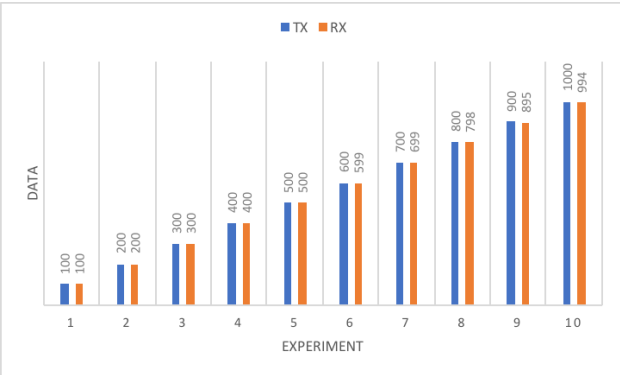


Fig. 10. Sending Data from IoT Devices to IoT Gateway

Figure 10 shows the sending of data from the IoT Device to the IoT Gateway by testing as many as 1000 data sent can be received as many as 994 pieces of data. After the data is received at the IoT Gateway, the data will be forwarded to the IoT Cloud. Figure 11 shows the display of the IoT Cloud.

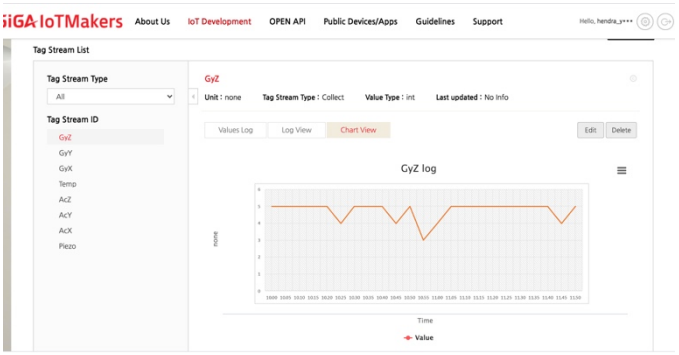


Fig. 11. IoT Cloud

Shown in figure 11 is a display of the IoT Cloud used. There is a graphic to the right of the layer. The graph is derived from data sent by the IoT Device to the IoT Gateway and forwarded to the IoT Cloud. The data contained in the IoT Cloud is the same as the data contained in the IoT Gateway as the Big Data Cluster.

IV. Conclusion

Jig detection system using this scanning method has a performance of 99.4%. Taking into account the reading distance and the distance between ultrasonic sensors, the IoT Device has a reading speed of 50µs for one reading cycle. Design a system that uses IoT Gateway as a Big Data Cluster so that the data can be ascertained properly recorded on the factory side. Whereas the main Big Data, the IoT Cloud has a performance of 100%, which is a backup of the IoT Gateway. With these results, the learning process at the Smart Learning Factory will be even better.

References

[1] M. Postranecky and M. Svitek, “Smart city near to 4.0 - An adoption of industry 4.0 conceptual model,” in *2017 Smart Cities Symposium Prague, SCSP 2017 - IEEE Proceedings*, 2017, pp. 15–19.

[2] M. Lom, O. Pribyl, and M. Svitek, “Industry 4.0 as a part of smart cities BT - 2016 Smart Cities Symposium Prague, SCSP 2016, May 26, 2016 - May 27, 2016,” in *Smart Cities Symposium Prague*, 2016, pp. 2–7.

[3] S. Kumar, P. Tiwari, and M. Zymbler, “Internet of Things is a revolutionary approach for future technology enhancement: a review,” *J. Big Data*, vol. 6, no. 1, pp. 1–21, 2019.

[4] M. Taufik, A. E. Rakhmania, and Y. N. Afnani, “Prepaid water meter card based on internet of things,” in *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 732, no. 1, pp. 1–7.

[5] S. Schneider, “The industrial internet of things (IIoT): Applications and taxonomy,” in *Internet of Things and Data Analytics Handbook*, 2017, pp. 41–81.

[6] M. Kurth, C. Schleyer, and D. Feuser, “Smart factory and education: An integrated automation concept,” in *Proceedings of the 2016 IEEE 11th Conference on Industrial Electronics and Applications, ICIEA 2016*, 2016, pp. 1057–1061.

[7] D. C. Huang, C. F. Lin, C. Y. Chen, and J. R. Sze, “The Internet technology for defect detection system with deep learning method in

smart factory," in *2018 4th International Conference on Information Management, ICIM 2018*, 2018, pp. 98–102.

[8] H. Jo, S. Kang, H. J. Kwon, and J. D. Lee, "In-door location-based smart factory cloud platform supporting device-to-device self-collaboration," in *2017 IEEE International Conference on Big Data and Smart Computing, BigComp 2017*, 2017, pp. 348–351.

[9] D. J. Ahn and J. Jeong, "A Bigdata Search Engine Based Cloud Computing Network Architecture in Smart Factory Environment," in *Proceedings - 2019 2nd International Conference on Intelligent Autonomous Systems, ICoIAS 2019*, 2019, pp. 91–95.

[10] C. Nagpal, P. K. Upadhyay, S. Shahzeb Hussain, A. C. Bimal, and S. Jain, "IIoT Based Smart Factory 4.0 over the Cloud," in *Proceedings of 2019 International Conference on Computational Intelligence and Knowledge Economy, ICCIKE 2019*, 2019, pp. 668–673.

[11] C. S. Kim and S. B. Son, "A Study on Big Data Cluster in Smart Factory using Raspberry-Pi," in *Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018*, 2019, pp. 5360–5362.

[12] A. Narkglom, E. Boonyapalanant, and P. Koseeyaporn, "Design of Training Media for Internet of Things Training Based on Project-based Learning: A Case Study of Smart Factory Industry," in *Proceedings of the 2019 International Conference on Power, Energy and Innovations, ICPEI 2019*, 2019, pp. 118–121.

[13] A. Rajendra and P. P. Tasgaonkar, "Object recognition using horizontal array of ultrasonic sensors," in *2016 International Conference on Communication and Signal Processing (ICCSP)*, 2016, pp. 983–986.

[14] P. Newman, "THE INTERNET OF THINGS 2020: Here's what over 400 IoT decision-makers say about the future of enterprise connectivity and how IoT companies can use it to grow revenue," *Business Insider*, pp. 1–6, 2020.

A Case of Information Technology Value Competency Left Behind? A Scoping Review

Warren Gabryk
Department of Informatics
University of Pretoria
South Africa
wgabryk@gmail.com

Rennie Naidoo
Department of Informatics
University of Pretoria
South Africa
rennie.naidoo@up.ac.za

***Abstract* — With Decision Support Systems (DSS) being critical in the delivery of value throughout the business, specialized IT practitioners are required to be fully enabled with a business value competency to ensure this value is efficiently harnessed. In this research we seek to map the current state of IT business value research through a scoping review to understand the state of business value competency of IT practitioners. Although more than 258 articles in leading IT journals have value as the subject matter, the problem is that over the past 20 years only 14 research articles have had the subject of competency in IT with a focus placed on the value of IT. Analysis is mapped to a value co-creation process framework derived from a thematic analysis of the reviewed sources. This research contributes by highlighting current opportunities in IT value competency research and proposes a value co-creation process and the skillful performance that is required of IT practitioners to better meet the value expectations of business stakeholders.**

Keywords— Business Value, Competency, IT Business Value Competency, Decision Support Systems.

I. INTRODUCTION

Information Technology (IT) has considerable consequences, both constructive and possibly dire, for the value of a company [8], [25]. In recognition of this fact, businesses are driving substantial efforts to better harness the business value of IT in both tangible and intangible measures alike [17]. The effort to harness the business value of IT can be evidently observed by the expected IT spend globally, meeting $3.4 trillion in 2020 and decision intelligence to be in 33% of large organizations by 2023 [14], [53]. Despite the mammoth investment and the pursuit of value standing at the forefront, competencies of IT professionals are being questioned [15], [25]. Given the lack of competency research conducted to enable IT practitioners, it is unclear whether IT practitioners are sufficiently equipped to lead the charge in delivering business value through the DSS medium as best as possible.

Before the point of competency in value is addressed, the concept of value should be outlined. Linguistically, value is "the regard that something is held to deserve; the importance, worth, or usefulness of something" [35]. The linguistic definition of value has not brought about certainty as to what value exactly is. To further muddle the case of value within the business environment, multiple interpretations lead to differing expectations of what should be achieved in pursuit of value by the divergent set of stakeholders within business [3], [7], [16], [25], [34], [39], [45], [47]. Accounting professionals will measure value based on accounting standards, such as IFRS13, business operational professionals will measure value based on productivity increases, customer services professionals will measure value on customer satisfaction and senior management will value investment outcomes as examples [3], [22], [41]. None of the value measures are incorrect but are rather a product of their own environment. Every activity will appear as a nail needing to be hammered if you are an expert hammer user [33].

Value within the business environment is in the eye of the beholder and with all these perspectives having some reasonable justification; the application of an IT value process must be consensus driven. This justification, for a consensus-driven IT value process, emanates from the thinking that a co-creative approach that ensures full participation will allow for the creation of unique value opportunities [40]. For the purposes of this paper, a strict definition of value will not be adopted, but rather agreed that value can be expressed in an economic, financial or accounting manner, that it can be tangible or intangible in nature and can either be observed in a short period of time or over an extended duration and from multiple perspectives.

Our research makes the following contributions: First, in conducting a scoping review, this research highlights the current deficiency in competency-based research within the IT field in relation to deriving value. Research marked with core subject of value since 2000 in the selected journals amounts to 258 research articles, while only 14 IT value competency-focused research articles were uncovered. The analysis of research papers, analysed through the search process, suggests that, albeit progress is being made with regards to researching value in the IT environment, a mechanism to instil this in the business environment through competency is being left behind. As a by-product of this result, this research draws attention to competency for IT value as a research area that deserves increased attention.

Second, we derive a Business Value Competency Framework that integrates work done on value with Bassellier's IT Business Competency Model [6]. By the identification of practical competencies that should supplement IT practitioners, as well as a definition of a Value Process, a competence seeking, knowledge worker aligned IT practitioner will be armed with the tools in order to guide value creation within the business environment through the criticality of DSS. The framework maps the competencies domain areas and associated competencies that are required to successfully identify, propose

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

and realise value within the business environment that have been addressed in the literature or which remains a knowledge gap. Our research question was: What is known from the existing IT literature about the competencies that are required of IT practitioners to create business value from IT? Given the level of research of competency completed in the field of IT, insight into how much of this is focused on the value in IT will highlight what efforts can be taken to develop the business value of IT as an individual and core competency.

DSS specialists bring unique understanding of the business decisions making process and hold technical knowledge to harness the power of IT and are expected to deliver value throughout the business through the implementation of DSS [46]. DSS specialists will need to succinctly rise to the occasion to understand value, IT value and the business value to successfully justify system investment and sustaining value using DSS within business. Exploring research efforts to solve IT business value problems will highlight areas where a competency has been successful in support of problem resolution, offer insight into best practices and enabling IT professionals to deliver beyond historical achievements.

II. Conceptual Foundations

Even though competency was first identified as a concept in the 1950's, competency and the modelling thereof rose to the forefront in the early 1970's [26]. In the last 50 years, a single view on competency is still to be set [26], [28]. Divergent views on competency in research are; a measurable human characteristic required for high performance, or a firm's ability to utilise resource through process execution to attain a target, [2], [32]. Businesses cannot afford to have business-illiterate IT practitioners. Specific competency enables business organisations to attract specialist staff, deliver concentrated efforts and effectively make use of their IT investments in support of operations and in alignment with strategy [7]. While research has been conducted into an IT practitioner's competency relating to knowledge management, firm performance, innovation, business managers' IT competency, business acumen of IT practitioners and core technical skills, competency regarding IT value has gone unaddressed [5], [6], [12], [19]. Without a competency which specifically focuses on the value of IT, as by way of understanding value and a collaborative value process that can govern value creation, the opportunity to prevent value-destruction and promote value-construction for businesses may be missed. For the purposes of this research competency has been adopted as a set of defined characteristics that ensure skilful performance of a business objective or task, ether individually or as a collective process.

Val IT, a business applicable framework, provides a governance-based, process-orientated framework to guide management and executives to deliver value from IT investment [54]. While Val IT has offered insight into processes, frameworks and collaborative requirements to deliver value from IT, it remains a rulebook and set of guiding principles only [31]. Although frameworks and process rulebooks, such as Val IT, make inroads for the delivery of value from IT, these approaches only enable the earlier stages of the five-stage model of expertise developed by Dreyfus and Dreyfus [13], [44]. The earlier stages of the Dreyfus and Dreyfus's five-stage model focus on the development of rulebooks and skilful enactment of these defined approaches, while the latter focus on the improved mastery and further development of these approaches [44]. Val IT has developed tremendous inroads into enabling businesses to deliver IT, however, there exists opportunity for IT value competency-based approaches to be developed in order to enable business not to just follow a governance framework but to master, beyond the point of following, to that of leading and developing within itself.

While issues surrounding IT value are continually being investigated in the IT field, the situation is clear with regards to the IT industry: IT practitioners, business practitioners and financial practitioners do not have or require a formally defined competency in order to articulate the value of IT. IS practitioners are also not enabled to understand the financial value or financial statements, nor are accounting practitioners knowledgeable enough in the field of IT to articulate its value. It comes to follow that the purpose of this research is to explore the current landscape of competency as applied to the IT discipline. Specific focus is placed on the application of competency applied to deriving value from IT investments. The research is primarily positioned to highlight the current research published with regards to competency and IT value and the direction for further research efforts will also be proposed. In support of further research endeavours, associated competency gaps will be added to Bassellier's IT Business Competency Model to propose an IT Business Value Competency for IT Practitioners.

III. Research Approach

The approach adopted for this research is a scoping review study [1], [37]. A scoping review approach was selected for this research to concentre on mapping published research papers to an IT Business value competency. In alignment with a consistent scoping review approach [1], [27], the following stages have been charted: research question, search strategy, research selection, data charting and results. With the research questions already being addressed in the introduction, this section starts with the search strategy stage of the scoping review.

A. Research Startegy

In execution of the search process, 19 IT journals have been selected to be part of this research. While it is understandable that the selected academic journals are not collectively exhaustive of all information systems and information technology knowledge, they may be considered the thought leaders and primary driving influencers in the IT discipline.

TABLE I.

Selected Journals	
Journal of the Association for Information Systems	Management Information Systems Quarterly
Communications of the Association for Information Systems	Scandinavian Journal of Information Systems
AIS Transactions on Human-Computer Interaction	European Journal of Information Systems
AIS Transactions on Replication Research	Information Systems Journal
Pacific Asia Journal of the Association for Information Systems	Information Systems Research

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Selected Journals	
MIS Quarterly Executive	Journal of Information Technology
Business and Information Systems Engineering	Journal of Management Information Systems
Journal of Information Systems and Technology Management	Journal of Strategic Information Systems
Journal of Information Technology Theory and Application	The Journal of the Southern Association for Information Systems
Journal of the Midwest Association for Information Systems	

Each Journal identified was queried for value as a concept as well as competency as a concept. The way the query process was conducted for each of these journals was twofold: firstly, an independent academic library database was queried using keywords of value and competency together and, secondly, the research specific search database was used to search for value and competency as subjects, where applicable.The search strategy executed has been done with the time frame boundary limit being placed.

B. *Research Selection*

The selection of findings relating to the research selection review are broken up into 2 parts: in the first part a quantitative selection of the research efforts being conducted into competency, value and value competency within the IT environment, and the second part concentrating on qualitative selection of the value-competency having been researched since 2000. Regarding the quantitative findings of the search results, Table 2 is used to compare the amount of research conducted within the boundary limits of this scoping review. When comparing the amount of research conducted between 'Value' and 'Value Competency' on a 'Keyword' search basis, the result supports a case of competency left behind. Only 338 research articles have concentrated on Competencies relating to Value over the past 20 years in comparison to 541 and 1237 for Competency and Value based research, respectively.

TABLE II.

Selected Journals Search Results		
Concept	***Search Area***	***Count***
Competency	Keyword Search	541
	Subject Search	54
Value	Keyword Search	1237
	Subject Search	258
Value AND Competency	Keyword Search	338
	Subject Search	3

Research selection involved both data collection and data analysis to deem what research is applicable to the research objective, purpose and questions posed thus far. Starting from the subject of competency being strictly defined as the departure point and then analysing if the concept of business value has been addressed, the following data was extracted from each article and collated into data spreadsheets to allow for an easier consolidation, analysis and applicability assessment:

- Philosophical and Methodological classification of the research article.
- Competency based theory applied (if applicable).
- The research questions of the research article.
- Summary of the findings and further research suggestions of the research article.

Having collected the above-mentioned data, it is analysed and compared to using 4 categories: value, business competency, value domain competency and value process. Value concentrates on whether the value (being built on by the research through competences under focus) is tangible or intangible. The definitions for 'tangible' and 'intangible' used in the data analysis are [21]:

- Tangible (T): physical value such as cash, revenue, capital equipment, inventory and investment bonds.
- Intangible (I): non-physical value such as internal business processes, decision support, team experience, trademarks and patents.

A business value competency categorisation is used to analyse the research collected from a competency perspective. The analysis is conducted by assessing if the sourced research article investigates a phenomenon that can be classified into business competences as defined by Bassallier [6]:

- Organisational (O) business competency is focused on providing a holistic, interconnected and comprehensive perspective on the business in which IT is an enabler.
- Interpersonal and Management (IM) business competency is focused on the competencies of the IS practitioner with regards to leadership, communication and social networking skills in support of engagement within the business environment.

A value competency domain classification is used to analyse the research collected from a value competency domain perspective. With economic, accounting and finance domains being highlighted as domains which hold value and business value as their core, these three are applied in this classification. The analysis is conducted by assessing if the sourced research article investigates a phenomenon employing fundamentals and/or characteristics from competencies that holds value at its core. In this regard:

- The Economic (EC) competency domain relates to the technical domain that is involved with expressing the social science discipline that focused on investigations of the transformation of scarce resources into wealth and value. To justify value economically would include justifications in pursuit of marginal gain of an economic eco-system.
- Financial (FC) competency is associated with the technical domain that is involved with stating the impact IT holds on business' market investments and cash generation mechanisms and the way IT influences the financial environment.
- Accounting (AC) competency is coupled with the professional domain that is involved with voicing the frameworks, principles and rules associated with the

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

process of recording, reporting and presenting the financial records of the business environment.

A value process categorisation is used to analyse the research collected against its applicability to the proposed Value Process. The analysis is conducted by assessing if the sourced research article investigates a phenomenon in a manner with a focus that would be applicable to the objectives, activities and associated competencies of the Value Process. To this end:

- Value Discovery (VD) relates to activities and supporting competencies that are focused on the identification of value opportunities within the business.
- Value Proposition (VP) relates to activities and supporting competencies that are focused on value case development and preparation of the solutions to add value.
- Value Realisation (VR) relates to activities and supporting competencies that are focused on ensuring that the value is delivered and accounted for.

C. Data Charting

When comparing the amount of research conducted between Value and Value Competency on a 'Subject' search basis, the result supports a case of competency left behind. Only 14 research articles have concentrated on Competencies relating to Value over the past 20 years in comparison to 54 and 258 for 'Competency' and 'Value' based research, respectively. As noted above, starting from a keyword search of competency being strictly defined as the departure point and applying a thematic analysis of the concept of 'business value' being addressed, the information in Table 3 tabulates a comparative analysis of research found with relevance to the concepts defined.

TABLE III.

Scoping Review Analysis - Value Co-creation Process Framework										
Article	***Value***		***Business Competence***		***Value Competency Domains***			***Value Process***		
	T	I	O	IM	EC	FC	AC	VD	VP	VR
[30]	X				X	X			X	X
[4]		X	X		X				X	
[29]		X	X		X				X	
[48]	X	X	X	X	X			X	X	X
[25]	X	X	X	X	X		X	X	X	X
[22]		X		X						X
[42]		X		X				X	X	
[8]	X		X			X				X
[52]	X		X	X	X			X	X	X
[50]		X		X					X	
[6]		X	X	X				X	X	X
[51]		X	X	X				X	X	X
[38]		X	X	X				X	X	X
[43]	X	X			X				X	X

A deviation from protocol has been included in the data-charting and results-scoping review. Taking into consideration that Bassallier's theory of IT Business Competency has been the guiding theory for investigation in this research, all competency and value-based academic work that references Bassallier's theory of IT Business Competency has been included for analysis post initial search strategy [6].

IV. Results

The following is observed from the selected journals highlighted in Table 3 above:

- Most of the competency-based research within IT focuses on intangible value rather than tangible value.
- Bassallier's business competencies for IT practitioners are the most prominent within IT research that focuses on value, when compared to financial, economic, and accounting based competencies.
- In all instances, where research focuses on tangible value, either economic, financial or accounting domain-based competencies were made use of in order to address the tangible value that the IT initiative sought to deliver.
- When tangible value is being addressed or used in order to justify the value of IT, value domain competencies are used as reference. On the other hand, when intangible value is addressed, most of the justification and alignment of value is that of Bassallier's theory of IT Business Competency [6].
- Social and influential competencies are noteworthy competencies which can influence majority spheres of the value process, value discussions within the business and strongly underpin the intangible value created.
- Analytical and discrete value-based domain disciplines underpin the tangible value created within business.

The research highlights a fragmented view of value competency, with no single item of research touching on the overarching process for value. While research gleans over identification, proposition, and measurement of value, these are posed as commentaries relating to a specific situation.

V. Discussion

Creating value through IT from within the business environment is no easy task. IT's value can be ascribed different value or meaning by different stakeholders in the business (depending on the role they play in the business) and if IT is not given proper value, or if not given correct value, this can be detrimental to business and an opportunity lost. In order to move towards a better way of describing the value of IT is fundamental to understand the meaning the various stakeholders give to IT value, how these stakeholders are trained and/or what competencies they hold within their respective disciplines, which process/es can be harnessed to facilitate value and how, through supplementing Bassallier's theory of IT Business Competency, is it possible to map the IT business value competencies that can help identify, create and/or realise the value of IT with greater efficiency [6].

- In pursuit of arriving at a knowledge-based view of value and increasing value from within the business

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

environment, this discussion focuses on the following points:

- Value: what is value within the context of IT and further enablement in a fluid business environment?
- Value Process: a Value Process that makes use of specialist knowledge workers.
- Value Process Competencies: competencies that can be established and driven in an IT Business Value Model to enable IS Practitioners.
- Application of Value Process, associated activities and competencies – application of competencies through the value process and in unison with existing IT and business process.

To better understand what the concept of value within the business environment is, it is necessary to first understand what value means. Due consideration is given to the fact that value holds different meanings and interpretations to the multiple stakeholders within the business environment. An import facet is for an IT business value competency to acknowledge the fact that value is simply that – fluid and dependant on context. In the face of this difference, value will still reside with the classification. Value can either display a tangible or intangible characteristic and, within each of those, can either be perpetual or temporary in nature. This value classification approach is proposed as a starting point for a divergent stakeholder viewpoint to engage in a value discussion. It is important that consensus towards an idea of value is worked on in a collaborative manner and moving away from a bipartisan view of who is correct and incorrect when value is derived. Through the review and analysis of the research conducted, specific emphasis on time horizons of value are not explicitly noted. This is the reason for its inclusion into the value classification of the Value Process. With it marked as a departure point for discussion, the time horizon of value is expected to gain additional focus.

In the vein of moving away from a bipartisan view of value, this research proposes a competency-based view of the value process that is defined by 3 distinct stages: value discovery, value proposition and value realisation. The value discovery stage is characterised by the identification and creative scenario synthesis of opportunities within the business environment that hold potential value. The value proposition stage is associated to the value case development and articulation of value ex-ante that when implemented will hold on the business. The final value process stage is associated with the realisation of the value. The value-realisation step of the value process is to ensure that the implementation and fulfilment of the activities associated to the opportunity are successfully delivered in the business, aligned to the value proposition and measured to ensure that what was discovered, proposed and implemented fulfilled obligations. The Value Process is sequentially liner from value discovery, value proposition through to value realisation. However, with due consideration of the creative and continuously evolving state of business and ITs influence within it, a recursive process route has been designed. The recursive process route allows for value discovery to be initiated at any time throughout the value process. As can be expected, while working within the value process, opportunities to discover and initiate concurrent value opportunities will constantly reoccur. It is expected that the proposed value process should naturally reoccur to infinity as the possibility of value discovery should be a non-ending cyclical process.

Having defined the value process, an understanding of what activities and competencies underpin this process is important to unpack. Just as other domains that are founded on with notion of value (accounting, economics and finance) at their core are building competency in IT competencies, so too is it imperative for IT practitioners to hold competencies in other value domain areas [10]. With value discovery concentrating on highlighting value opportunities, value proposition building the value case and value realisation relating to the implementation and measurement of the value resultant, the required competencies to conduct the value process are currently noted in literature, albeit fragmented. Value discovery, value proposition and value realisation would naturally include competencies that are identified by Bassallier's IT business competency model of organisation business competencies and interpersonal and management competencies. This would afford the IS practitioner competencies in being able to engage within their functioning of IT and engage in the functioning of departments that focus on value as a core discipline. It comes to pass that, deductively, injecting value competencies into the business from other disciplines would increase the effectiveness of value [7]. However, ensuring the domain from which value builds part of its core in the business environment is instrumental in ensuring a robust and sound value process competency. An elaborative list of competencies aligned to value process and associated value process activities is highlighted in Table 4 below.

As depicted in Table 4, there are 12 unique competencies that will be required throughout the Value Process and its associated activities. 2 of 12 noted competencies have been previously specified by Bassallier's theory of IT Business Competency and the remaining 10 stems from literature found in financial, economic, accounting and IT Value based domains. While gaps still exist in literature, the authors have highlighted the need and focus for each of the respective Value Processes and associated activities and competencies. This co-creation underpinned value process moves value solely away from the functional disciplines in a business towards that of a process-orientated, collaborative, knowledge worker approach. This makes inroads into better enhancing value discovery, proposition and realisation throughout the business and not leaving this up to a single function within the business to drive [40]. This approach will support greater alignment between business functions. Having addressed the fragmented literature of Value competencies within the IT environment, the proposal of a value process with associated activities and competencies that underpin the objective of creating a knowledge-based view of value within the business, this research would purpose the below IT Business Value competency model for IT Scholars and IT Practitioners. The supplements map and highlight the numerous gaps identified in the literature for future research.

IT Practitioners can also make effective use of these findings. As discussed thus far, this mapping of the IT Business Value competency affords the IS Practitioner the ability to lead a process-focused competency-based view of value from within

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

the business. It is not expected that the competencies and knowledge areas identified here are to be held of that of the foremost expert. Rather, armed with a general understanding, the leader of such process can leverage the knowledge and expert competencies in a business with the objective of value co-creation. It must be noted that while the IT Business Value Competence Model for IT Practitioners is IT specific in its current form, the model can be utilitarian model for uses and will be specifically applicable to DSS experts.

TABLE IV.

Scoping Review Analysis - Value Co-creation Process Framework			
Value Process	***Activity***	***Competency***	***Examples***
Value Discovery	Business Assessment	Research	[26]
	Opportunities Identification	Innovation	[26], [52]
	Hypothesis Generation	Analytical	[26]
	Project Drivers Consolidation	Advisory	[26]
	Solution Vision	Strategic	[11], [39]
Value Proposition	Solution Definition	Technical	[9], [26]
	Business Case Development	Writing	[26], [35]
	Economic Impact Articulation	Analytical	[11]
	Finance Impact Articulation	Regulatory Governance	[27]
	Accounting Record Articulation	Analytical	[10]
	Present Business Solution	Communi-cation	[15], [26]
	Benefits Articulation	Influence	[15], [26]
	Attain Project Approval	Influence	[26], [49]
Value Realisation	Solution Establishment	Portfolio Management	[26]
	Solution Tracking	Leadership	[26]
	Goal Alignment Maintenance	Influence	[26]
	Benefits Realisation	Influence	[2], [26]

CONCLUSION

With DSS being crucial to driving the value of business, understanding the business value competency of IT professional is critical in further enabling success. To this end the objective of this scoping review identify areas where this competency can be further developed within the IT profession. Although more than 258 articles in leading IT journals addressed the value problem over the past 20 years, only 14 research articles have been uncovered which address competency in IT with a focus on the business value of IT. Of these 14 articles, most of the findings pertaining to IT business value competency emphasise the social and behavioural aspects of business competency, such as organisational awareness and interpersonal and managerial qualities. We present a mapping of areas addressed and research gaps that need attention, structured according to a value co-creation process framework. The majority of value-based and competency-based research in the IT field concentrates on economic aspects of value within the business. Our mapping suggests that there exists an opportunity for future research efforts to focus on financial and accounting-based value of IT within the business. Considering that business focuses on money generation, finance, statutory reporting and accounting, it would be prudent to equip IT practitioners with these value-centric competencies. Given that the business value of IT is based on the varying perceptions of business stakeholders and custodians of IT governance, enabling value-based competencies that focuses on using data from economic, financial and accounting practices will move IT value from perception-based approaches to more evidence-based approaches. Given the research contribution of highlighting opportunities in IT value competency research and the introduction of a value co-creation process together with the skilful performance that is required of IT practitioners to deliver value, we trust that further academic and industry-based research is conducted in the future in order to drive the value of DSS through an evidence-based business value IT competency model for IT practitioners.

VI. REFERENCES

[1] H. Arksey and L. O'Malley, "Scoping Studies: Towards a Methodological Framework," *International Journal of Social Research Methodology*, Volume 8, pp. 19-32, February 2005.

[2] C. Ashurst, N.F. Doherty and J. Peppard, "Improving the Impact of IT Development Projects: The Benefits Realization Capability Model," *European Journal of Information Systems*, Volume 17, pp. 352-370, August 2008.

[3] S. Athey and M. Luca, "Economists (and Economics) in Tech Companies," *Journal of Economic Perspectives*, Volume. 33, pp. 209–230, Wnter 2019.

[4] J. Baker, D. Jones, Q. Cao and J. Song, "Conceptualizing the Dynamic Strategic Alignment Competency," *Journal of the Association for Information Systems*, Volume 12, pp. 299 – 322, April 2011.

[5] G. Bassellier, 2014. "Business Competence and Acumen of Information Technology Professionals: Conceptual Development and Influence on IT-Business Partnerships," *MIS Quarterly*, Volume 28, pp. 302–312, January 2014.

[6] G. Bassellier and I. Banbasat, 2004. "Business Competence of Technology Professionals: Conceptual Development and Influence on IT-Business Partnerships," *MIS Quarterly*, Volume 28, pp. 673–694, December 2014.

[7] G. Bassellier, B.H. Reich and I. Benbasat, "Information Technology Competence of Business Managers: A Definition and Research Model," *Journal of Management Information Systems*, Volume 17, pp. 159–182, April 2001.

[8] M. Benaroch and A. Chernobai, "Operational IT Failures, IT Value-Destruction, and Board-Level IT Governance Changes," *MIS Quarterly*, Volume 41, pp. 729 – 762, September 2017.

[9] E. Berghout and C.W. Tan, "Understanding the Impact of Business cases on IT investment decisions: An analysis of municipal e-government projects," *Information and Management*, Volume 50, pp. 489-506, November 2013.

[10] P.C. Brewer, J.E. Sorensen and D.E. Stout, "The Future of Accounting Education: Addressing the Competency Crisis," *Strategic Finance*, Volume 96, pp. 29–37, August 2014.

[11] B. Carlsson and G. Eliasson, "The Nature and Importance of Economic Competence," *Industrial and Corporate Change*, Volume 3, pp. 687-711, March 1994.

[12] J.P. Downey and S. Zeltmann, "The Role of Competence Level in the Self-Efficacy-Skills Relationship: An Empirical Examination of the Skill Acquisition Process and its Implications for Information Technology Training," *International Journal of Training and Development*, Volume 13, pp. 96–110, June 2009.

[13] H.L. Dreyfus and S.E. Dreyfus, *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*, New York: Free Press, 1986.

[14] Gartner. (2020, May. 13) *Global IT Spending Press Release* [Online], Available: https://www.gartner.com/en/newsroom/press-releases/2020-05-13-gartner-says-global-it-spending-to-decline-8-percent-in-2020-due-to-impact-of-covid19

[15] M. Ghasemaghaei, "Does data analytics use improve firm decision making quality? The role of knowledge sharing and data analytics competency," *Decision Support Systems*, Volume 120, pp.14–24, March 2019.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

[16] S.M. Glover, M.H. Taylor and Y.J. Wu, "Current Practices and Challenges in Auditing Fair Value Measurements and Complex Estimates: Implications for Auditing Standards and the Academy," *Auditing*, Volume 36, pp. 63–84, February 2017.

[17] J.M. Goh and A.E. Arenas, "IT Value Creation in Public Sector: How IT-Enabled Capabilities Mitigate Tradeoffs in Public Organisations," *European Journal of Information Systems*, Volume 29, pp. 1-19, January 2020.

[18] M. Goldberg, A. Kieninger, G. Satzger and H. Fromm, "Retained Organizations in IT Outsourcing: Linking Organization Design to Outsourcing Management Problems," *Business and Information Systems Engineering*, Volume 59, pp. 111–124, April 2017.

[19] S.R. Gordon and M. Tarafdar, "How do a Company's Information Technology Competences Influence its Ability to Innovate?," *Journal of Enterprise Information Management*, Volume 20, pp. 271–290, 2007.

[20] A. Heirman and B. Clarysse, "Do Intangible Assets at Start-Up Matter for Innovation Speed?" *Frontiers of Entrepreneurship Research*, Volume 24, pp. 303 – 315, 2004.

[21] S. Henningsson and C. Øhrgaard, "IT Consultants in Acquisition IT Integration," *Business and Information Systems Engineering*, Volume 58, pp. 193 – 212, May 2016.

[22] IASPlus. (2013, January). *IFRS13* [Online], Available: https://www.iasplus.com/en/standards/ifrs/ifrs13

[23] R. Jung and C. Lehrer, "Guidelines for Education in Business and Information Systems Engineering at Tertiary Institutions," *Business and Information Systems Engineering*, Volume 59, pp. 189-203, June 2017.

[24] P.G. Keen, "Value Analysis: Justifying Decision Support Systems," *MIS Quarterly*, Volume 5, pp. 1 – 15, March 1981.

[25] K. Kim, S. Mithas and M. Kimbrough, "Information Technology Investments and Firm Risk Across Industries: Evidence from the Bond Market," *MIS Quarterly*, Volume 41, pp. 1347–1367, December 2017.

[26] R. Klendauer, M. Berkovich, R. Gelvin, J.M. Leimeister and H. Krcmar, "Towards a Competency Model for Requirements Analysts," *Information Systems Journal*, Volume 22, pp. 475–503, November 2012.

[27] T. Kude, M. Lazic, A. Heinzl and A. Neff, "Achieving IT-Based Synergies Through Regulation-Oriented and Consensus-Oriented IT Governance Capabilities," *Information Systems Journal*, Volume 28, pp. 765–795, December 2018.

[28] F.D. Le Deist and J. Winterton, "What is competence?," *Human Resource Development International*, Volume 8, pp. 27–46, March 2005.

[29] O.K.D. Lee, P. Xu, J.P. Kuilboer and N. Ashrafi, "Idiosyncratic Values of IT-Enabled Agility at the Operation and Strategic Levels," *Communications of the Association for Information Systems*, Volume 39, pp. 242 – 266, 2016.

[30] R. Legoux, P.M. Leger, J. Robert and M. Boyer, "Confirmation Biases in the Financial Analysis of IT Investments," *Journal of the Association for Information Systems*, Volume 15, pp.33 – 52, February 2014.

[31] R. Lombardi, M. Del Giudice, A. Caputo, F. Evangelista and G. Russo, "Governance and Assessment Insights in Information Technology: the Val IT Model," *Journal of the Knowledge Economy*, Volume 7, pp. 292–308, March 2015.

[32] A.F. Marrelli, J. Tondora and M.A. Hoge, "Strategies for Developing Competency Models," *Administration and Policy in Mental Health* Volume 32, pp. 533–561, May 2005.

[33] A.H. Maslow, *The Psychology of Science: A Reconnaissance*, Harper and Row, 1966.

[34] B.N. Melville and K. Kraemer, "Information Technology and Organizational Performance: Integrative Model of IT Business Value," *MIS Quarterly*, Volume 28, pp. 283–322, June 2004.

[35] P.A. Nielsen and J.S. Persson, "Useful Business Cases: Value Creation in IS Projects," *European Journal of Information Systems*, Volume 26, pp. 66–83, January 2017.

[36] Oxford Dictionary (2020) *Value* [Online], Available: https://www.lexico.com/en/definition/value

[37] G. Paré, M.C. Trudel, M. Jaana and S. Kitsiou, "Synthesizing Information Systems Knowledge: A Typology of Literature Reviews," *Information and Management*, Volume 52, pp. 183-199, March 2015.

[38] J. Peppard, R. Lambert and C. Edwards, "Whose Job is it Anyway?: Organizational Information Competencies for Value Creation," *Information Systems Journal*, Volume 10, pp. 291–322, October 2000.

[39] J. Peppard and J. Ward, "Beyond Strategic Information Systems: Towards an IS capability," *Journal of Strategic Information Systems*, Volume 13, pp. 167–194, February 2004.

[40] C.K. Prahalad and V. Ramaswamy, "Co‐Creating Unique Value with Customers," *Strategy and Leadership*, Volume 32, pp. 4‐9, June 2004.

[41] T.J.W. Renkema, *The IT Value Quest: How to Capture the Business Value of IT-Based Infrastructure*, John Wiley and Sons Ltd, 2000.

[42] L.K. Roses, J.C.B. Brito and G.J.H. Lucena Filho, "Conversational Competences Model for Information Technology and Business Strategic Alignment," *Journal of Information Systems and Technology Management*, Volume 12, pp. 125-144, January 2015.

[43] K.A. Saeed, V. Grover and Y. Hwang, "The Relationship of e-Commerce Competence to Customer Value and Firm Performance: an Empirical Investigation," *Journal of Management Information Systems*, Volume 22, pp. 223–256, July 2005.

[44] J. Sandberg, L. Rouleau, A. Langley and H. Tsoukas, *Skilful Performance: Enacting Capabilities, Knowledge, Competence and Expertise in Organizations*, Oxford University Press, 2017.

[45] A. Saunders and E. Brynjolfsson, "Valuing Information Technology Related Intangible Assets," *MIS Quarterly*, Volume 40, pp. 83–110, March 2016.

[46] J.P. Shim, M. Warkentin, J.F. Courtney, D.J. Power, R. Sharda and C. Carlsson, "Past, present, and future of decision support technology," *Decision Support Systems*, Volume 33, pp. 111–126, June 2002.

[47] P.P. Tallon, K. Kraemer and V. Gurbaxani, "Executives Perceptions of the Business Value of Information Technology: A Process-Oriented Approach," *Journal of Management Information Systems*, Volume 16, pp. 145–173, April 2000.

[48] H. Topi, H. Karsten, S.A. Brown, J. Alvaro, B. Donnellan, J. Shen, B.C. Tan and M.F. Thouin, "MSIS 2016 Global Competency Model for Graduate Degree Programs in Information Systems," *Communications of the Association for Information Systems*, Volume 40, pp. I – 107, November 2017.

[49] C. Urquhart, "Analysts and Clients in Organisational Contexts: A Conversational Perspective," *Journal of Strategic Information Systems*, Volume 10, pp. 243–262, September 2001.

[50] A.H. Van de Ven, "Running in Packs to Develop Knowledge-Intensive Technologies," *MIS Quarterly*, Volume 29, pp. 365-378, June 2005.

[51] H.T. Wagner, D. Beimborn and T. Weitzel, "How Social Capital Among Information Technology and Business Units Drives Operational Alignment and IT Business Value," *Journal of Management Information Systems*, Volume 31, pp. 241-272, July 2014.

[52] G. Westerman and M. Curley, "Building IT-Enabled Innovation Capabilities at Intel," *MIS Quarterly Executive*, Volume 7, pp. 33 – 48, September 2008.

[53] Which-50. (2020, June. 22). *Data And Analytics Technology Trends For 2020* [Online], Available: https://which-50.com/gartner-identifies-top-10-data-and-analytics-technology-trends-for-2020/

[54] C. Wilkin, J. Campbell, S. Moore and W. van Grembergen, "Co-creating value from IT in a contracted public sector service environment: Perspectives on COBIT and Val IT," *Journal of Information Systems*, Volume 27, pp. 283–306, July 2013.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Secure Architecture for Inter-Healthcare Electronic Health Records Exchange

Oluwaseyi Ajayi
Department of Electrical Engineering, City College of New York,
New York, USA
Oluwaseyi.j.ajayi@gmail.com

Meryem Abouali
Department of Electrical Engineering, City College of New York,
New York, USA
Maboual000@citymail.cuny.edu

Tarek Saadawi
Department of Electrical Engineering, City College of New York,
New York, USA
sadaawi@ccny.cuny.edu

Abstract—**The increase in cyberattacks against the healthcare system, notably Electronic Health Records (EHRs) breaches, has cost the healthcare providers more in recent years. This situation is predicted to increase in the coming years as the healthcare systems are proposing a consortium EHRs repository. Due to this reason, it is crucial to deploy solutions that can ensure the security of shared health records. More specifically, maintaining the integrity and consistency of shared EHRs becomes pertinent. In this on-going research, we propose a blockchain-based solution that facilitates a scalable and secured inter-healthcare EHRs exchange. These healthcare systems maintain their records on individual private blockchain networks, and the blockchains interact to exchange patient health history based on request. The proposed solution verifies the integrity and consistency of requests and replies from other healthcare systems. It presents them in a standard format that can be easily understood by different healthcare nodes. The verification steps guard against malicious activities on both stored and in transit EHRs from insider and outsider threat actors. We evaluate the security analysis against frequently encounter outsider and insider threats within a healthcare system. The preliminary result shows that the architecture can detect and prevent threat actors from uploading compromising EHRs into the network and prevents unauthorized retrieval of patient's information.**

Keywords— Blockchain, Electronic Health Records, Compromised EHR, Healthcare System, Security, Data Integrity, Consistency.

I. Introduction

The tremendous increase in the healthcare record breaches has made healthcare interoperability more challenging to achieve. In 2019, around 572 recorded attacks in the U.S. healthcare industries resulted in over 41 million patient record breaches, and it is estimated to jump up by 60% in 2020 [1]. The effects of these attacks are estimated to cost the industry about $1.4 billion a year. Although ransomware attack accounts for about 58% of the total breach, staff members inside the healthcare organization were responsible for about 9.2% of the data breach in 2019 [2]. Due to the prevalence of attacks on patient records, there is an urgent need to deploy a more secure method to protect Electronic Health Records (EHRs) shared among different healthcare systems, especially now that healthcare systems are proposing more robust interoperability. For the proposed interoperability to work, the integrity and consistency of the stored or shared EHRs must be guaranteed. Thus, it is crucial to have a solution that detects and prevents malicious activities in EHRs. Over the years, healthcare systems have utilized a firewall to protect stored EHRs [3-7]. [3] iimplements the firewall to serve as an anomaly-based intrusion detection system (IDS). In the implementation, the firewall is either configured as a packet filtering firewall or status inspection firewall. The authors in [8] put forward encryption to ensure the security of EHRs during the exchange process. This approach was designed by Health Insurance Portability and Accountability Act (HIPAA) to secure EHRs when viewed by patients or when creating, receiving, maintaining, or transmitting Patient Health Information (PHI) by mobile devices. Despite the success of the approaches, the malicious intruders still find ways to subvert these protection systems and gain unauthorized access to EHRs.

Healthcare providers believe that their data is secured as far as it is encrypted. Although encryption guarantees the confidentiality of such data, consistency, and integrity are not guaranteed. [9] proposed a message authentication code algorithm (MAC) for detecting any changes in stored data. Although this approach detects changes in the stored data, it is not practical for extensive data because downloading and calculating MAC of large files is overwhelming and time-consuming. Another method described in [9] secures cloud data integrity by computing the hash values of every data in the cloud. This solution is lighter than the first approach in [9]. However, it requires more computation power, especially for massive data; hence, it is not practical. The authors in [10] employ the third party to coordinate activities of the database. The problem with this approach is that the data is vulnerable to man-in-the-middle or single-point-of-failure attack.

Further research has put forward the application of blockchain technology in handling, protecting, and interacting IoT devices with personal EHRs [10-17]. The approaches described in these researches prove effective in handling and protecting stored personal EHRs. However, the proposed solutions cannot be applied to the EHRs exchanged between two or more healthcare systems as they are primarily focusing on securing and protecting personal EHRs. In this research, we propose a solution that leverages the tamper-proof ability, data immutability, and distributive ledger ability of blockchain to share EHRs among different healthcare systems securely.

The new dimension in the healthcare industry is the interoperability of different healthcare systems. The interoperability is important because a patient's diagnosis

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

and treatment journey can take them from a physician's office to an imaging center or the operating room of a hospital. Each stop generates a record, such as doctor's notes, test results, medical device data, discharge summaries, or information essential to the social determinants of health, which become part of a patient's electronic health record in each setting. For the best outcome, this health information should be accessible and securely exchanged among all sources that accompany the patient's treatment every step of the way. This interoperability will strengthen care coordination and improve safety, quality, efficiency, and encouragement of robust health registries. However, most of the available solutions in hospitals use fax messages for EHR exchange between healthcare systems, and cloud database for storing EHRs. The significant problems facing the currently available solution are (i). The medium of exchange can be hacked, thereby compromising the integrity and consistency of the shared data. (ii) The database housing the EHRs can be hacked, and data can be manipulated or deleted. (iii) Lack of universal format for EHRs exchange makes it difficult to detect and prevent malicious activities.

We propose a solution that ensures the integrity and consistency of shared health data, presents a standard format for exchanging EHRs, and detects any malicious activities on stored and shared EHRs. Hence, the contributions of our work can be summarized as follows:

- We propose a blockchain-based architecture that facilitates a scalable and secured inter-healthcare EHRs exchange among different healthcare systems.
- The proposed architecture detects and prevents malicious activities on both stored and shared EHRs from either outsider or insider threats.
- The architecture verifies the integrity and consistency of EHR requests and replies, then presents them in a standard format easily understood by the different healthcare systems.
- The architecture permanently stores the verified EHRs in a distributive ledger and shares with other health care systems securely when requested.
- The proposed architecture is robust to a new healthcare system joining the network in real-time.

The remainder of this paper is organized as follows: related works on blockchain application in healthcare are discussed in Section II. Section III describes the proposed architecture, while the preliminary result is presented in section IV. Section V presents the conclusions of this paper and possible future works.

II. RELATED WORKS

A. *Blockchain application in healthcare*

First introduced as the technology behind bitcoin in 2008 [18], blockchain was implemented to solve the double-spending problem in a cryptocurrency called bitcoin. Since its inception, diverse areas have seen the application of blockchain technology. e.g., health system [10-17], data integrity [19], [29-30], intrusion detection system [20-22], and so on. Blockchain is an append-only public ledger that records all transactions that have occurred in the network. Every participant in a blockchain network is called nodes. The data in a blockchain is known as a transaction, and it is divided into blocks. Each block is dependent on the previous one (parent block). So, every block has a pointer to its parent block. Each transaction in the public ledger is verified by the consensus of most of the system's participants. Once the transaction is verified, it is impossible to mutate/erase the records [18]. Blockchain is broadly divided into two: public and private blockchain [23]. A public blockchain is a permissionless blockchain in which all nodes do verification and validation of transactions. e.g., Bitcoin, Ethereum. While private blockchains are permissioned blockchains where only nodes given permission can join and participate in the network. e.g., Hyperledger.

Blockchain application in EHRs is still in its inception. However, the potential it offers, the deficiencies and gaps it fills and ensuring the security and confidentiality of health data make it the forefront to be adopted in the healthcare industry nowadays. Blockchain technology has been applied to different kinds of researches to secure personal data. The authors in [24] propose a platform that enables a secure and private health record system by separating sensitive and non-sensitive data. The platform serves to share a patient's healthcare data with researchers without revealing the patient's privacy. The model successfully uses proxy re-encryption techniques to share a patient's sensitive data without revealing the patient's private key and adopting an asymmetric cryptography technique to encrypt these data while storing it on the cloud. Another similar work [25] proposes i-Blockchain, which uses a permissioned blockchain to preserve the privacy of the Patient's Health Data (PHD) and improve the individual's experience in data exchange. It allows only qualified individuals and Healthcare Service Providers (HSP) to join the network to prevent malicious attacks. It uses cold storage functions as off-blockchain storage and hot storage functions as the store where users temporarily put requested data in addition to a private key and a public key for secure data exchange.

Furthermore, [26] proposes the conceptual design for sharing personal continuous dynamic health data using blockchain technology. The approach is supplemented by cloud storage. The authors proposed using hash pointers to the storage location to solve the problem of sharing large-sized continuous-dynamic data while integrating blockchain and cloud storage. Extensive size data can be stored in an encrypted format on the cloud, and only the transactional data and metadata can be saved and shared on the blockchain. The authors in [27] propose a decentralized record management system (MedRec) to manage authentication, confidentiality, accountability, and data sharing of EHRs using blockchain technology. It is a modular design that integrates with patient's local data storage and encourages medical stakeholders to participate as miners. The result shows that the system enables the emergence of big data to empower researchers while engaging the patient and providers in the choice of release metadata. [28] proposes a new approach which joins blockchain and cloud computing network. In their work, they employ Amazon Web Services and Ethereum blockchain to facilitate the sematic level interoperability of EHRs systems without standardized data forms and formatting. The model proposes an interoperability data sharing framework that includes security through multilayer encryption, optical data

storage through Amazon Web Service, and transfer using the Ethereum blockchain.

Despite the researches on blockchain application in healthcare, most of the available solutions focus on securing and sharing personal EHRs, failing to address the security of health records shared between two or more healthcare systems. The lack of secure exchange of EHRs is the motivation for this work. The novelty in the proposed solution is to facilitate a scalable and secured inter-healthcare EHRs exchange while detecting and preventing malicious activities on the data. This novelty distinguishes our work from previous works.

III. THE PROPOSED ARCHITECTURE

The proposed architecture, which focuses on securing inter-healthcare EHRs exchange, is implemented on the Hyperledger Fabric blockchain platform. The Hyperledger fabric features a chaincode which houses the smart contract. The chaincode is deployed to the fabric, and each peer interacts with the code. The smart contract keeps the agreement among consortium members, and all participants run it. Fig. 1 shows a pictorial representation of the proposed architecture.

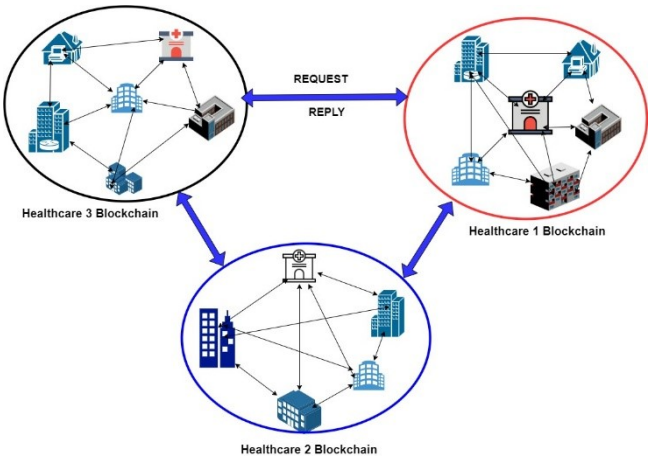


Fig. 2. The Proposed Architecture

The architecture comprises of different healthcare systems running private blockchain networks. Each blockchain network is independent of each other and features a unique smart contract that is written according to the healthcare system and HIPAA's policies. In each healthcare system, computers used in collecting patient records from peers in a private blockchain network. The peers in each network prepare, submit, verify, and keep a copy of the ledger of all transactions (i.e., patient's EHRs). The peers also run the consensus algorithm, thus validate transactions/blocks. The peers similarly validate the transactions, as described in our previous works [29,30]. However, unlike the previous works, which uses public-private blockchain networks, we set up a fully private blockchain network for each healthcare system.

In this current work, we focus on investigating a secured exchange of patients' records among different healthcare systems. In this implementation, the healthcare systems keep and maintain patient health information on individual private blockchain networks. We describe the process of interaction and exchange of the patients' EHRs among them. We deploy a smart contract on each private blockchain platform via chaincode. The smart contract handles the verifications of all transactions. In this context, we define a transaction as any of the following:

1. A patient's health information about to be stored into the blockchain network
2. A request for patients' medical history from another healthcare system.
3. A reply that carries the requested patient's information

This paper describes how the architecture achieves the verification and validation of requests and replies to mitigate malicious activities. (i.e., we describe how a request for a patient's health history undergoes different security verification and validation before being sent to the closed primary healthcare system.)

The proposed architecture is divided into three main steps, as shown below.

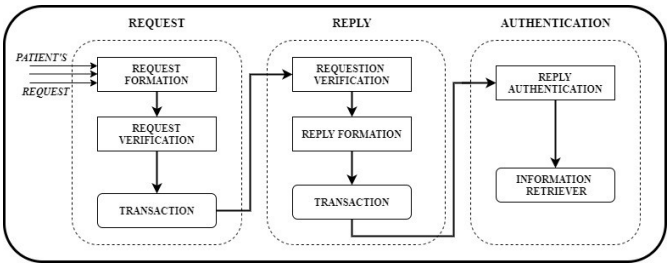


Fig. 3. Building blocks of the proposed architecture

A. Request

The request stage is subdivided into three categories: request formation, request verification, and transaction formation. The request is formed to obtain a patient's medical history from another healthcare center or medical test results from a laboratory. For example, a person who lives in New York, USA, travels to London, UK. If the person had to visit the hospital for treatment, the medical history should be retrieved from the New York hospital for better medical treatment. The medical records can be retrieved by preparing a request with information that is unique to the patient. During the process, a requester (doctor or nurse in the visiting hospital) supplies the required information to a developed script running on the peers. This script captures the patients' information such as name, date of birth, Social Security Number (SSN), name of the former healthcare system, and requester's unique code.

The script verifies the information and the identity of the requester. The request is developed into an agreed-upon format, digitally signed by the peer, and submitted as a transaction to the hospital's blockchain network. Apart from the submitted transaction, the peer (node) submits its information, which includes the requester's unique code, the MAC address, and the peer's transaction address. The smart contract run on the blockchain verifies the transaction's format, the requester code, and the peer's identities. The purpose of the verification is to detect and prevent all unauthorized activities on the transaction. Algorithm 1 below describes the snippet of the smart contract that

handles the verification process. Since the healthcare blockchains communicate via smart contracts, each smart contract running on the blockchain networks contains the processes described in algorithm 1.

The smart contract verifies the transaction by comparing the format with the stored transaction format. It checks the number of requests from the same peer within a minute and compares it to a maximum value. We set the maximum value of request from a peer to 1. It also checks if the destination address is found in the blockchain lookup table. It confirms the privilege of the peer to submit transactions by verifying its digital signature with its public key. It further checks for the peer's information in the information lookup table and verifies the health worker's code. Apart from the above, the smart contract controls each peer's right to assess patients' records. For a request to be successful, it must agree with a standard format, and the number of requests per minute must not exceed a maximum value. Also, the requester must be authorized, destination healthcare blockchain must be available and correct, peer's information must be correct, and the peer's public key must verify its private key. If any of these verification steps fail, the transaction is dropped. A successful request is validated and attached to the blockchain. The validation process has been described in our previous works [29,30].

Algorithm 1: Request Verification

Procedure: Verification **(Request)**
Inputs: formatted request/reply, Peer Information

```
If (Request agrees with Standard Format) and
    (No. of Requests/min <= Max) and
    (Requester code in authorized code sets) and
    (Destination in lookup table) and (Peer Information is verified)
     and (public key verifies private key):

        Validate Transaction
          Return Success
else:
          Return fail
          Drop transaction
end if
end procedure
```

B. *Reply*

After a successful validation process, the smart contract routes the request to the designated healthcare network based on the lookup table. The smart contract verifies the format of the received request and the requesting network. Algorithm 2 below describes the pieces of the smart contract that handles the verification of an incoming request. The source information, format of request, and digital signature of the source healthcare are verified. The supplied information about the patient is confirmed. For an incoming request to be successful, the request's format must agree with the standard, source information must pass verification step, requested EHRs must be available in the healthcare network, and the source public key must verify its private key. If any verification steps fail, the request is dropped, and a failed request is issued to the source network. A successful request is validated and attached to the blockchain. Based on the required information, the peers compete to prepare a reply by retrieving the patient's EHRs from the blockchain network. The peer that first prepares the information submits it as a transaction to the blockchain for verification. The transaction is verified as explained in algorithm1. A successfully verified reply is validated and routed back to the source network.

Algorithm 2: Reply Formation

Procedure: Reply formation by destination network
Inputs: request, Source Information (S.I).

```
If (Request agrees with Format) and
    (S.I in lookup table) and
    (Patient's EHRs in Destination Healthcare) and
    and (public key verifies private key):

          Return Success
          Validate the request and form reply
else:
          Return fail
          Drop request
end if
end procedure
```

C. *Authentication*

The source network verifies the incoming reply. The smart contract verifies the format of the information, the address of the reply's network, and the digital signature of the reply's network in a way showed in algorithm 3. When the verification process is successful, the transaction (reply) is validated and attached to the blockchain. The blockchain is updated, and the newly added block reflects on the ledger of every peer in the network. Every blockchain peer possesses a copy of this ledger. All blockchain peers receive the notification of the newly added block but do not have access to the block's content. The requesting node retrieves the information in the block, and a developed script converts it to a format that can easily be understood by the requester (healthcare work).

Algorithm 3: Reply Verification

Procedure: Verification **(Reply)**
Inputs: formatted reply, Source Information

```
If (Reply agrees with Standard Format) and
    (Request destination matches reply source) and
    (source Information in respective sets) and
    (source digital signature is verified):

          Validate reply
          Return Success
else:
          Return fail
          Drop reply
end if
end procedure
```

IV. RESULTS

We carry out the implementation of the proposed architecture in the laboratory. We set up two different Hyperledger fabric blockchain networks (I and II) with each network comprising of three peers. For each blockchain network, we installed Hyperledger Fabric *v2*, use *Golang v1.14* implementation for smart contract and *docker v19.0.3 and docker-compose v1.26.2, nodejs v12.18.1, npm v6.4.1*and *python v3.6*. All the blockchain peers run Linux Ubuntu *v18.04* operating system. The smart contract is written as described above and deployed into the blockchain channels. The MAC and transaction addresses of all peers are written as lists in the smart contract. A list that stores the requester code is also included in the smart contract. The format of requests and replies and the requester's unique codes are coded in the smart contract. Apart from this, the smart contract also contains a lookup table that stores information about the blockchain II. We write a similar smart contract (with blockchain I information) for blockchain II. A transaction (request for patients' medical history) was prepared as explained in section III and submitted as a transaction to the blockchain network I. In the security analysis, we anticipated and tested ways an insider or outsider malicious intruder could attempt to compromise a request within a healthcare system. We performed the attack and presented the result obtained from our architecture. In the outsider attack, we implement how the architecture detects and prevents unauthorized peers from submitting a transaction to the blockchain network. While for the insider attack, we showed how the architecture detects a compromised peer and prevents it from inserting its transaction.

A. SECURITY ANALYSIS

1) Outsider Threat Detection

We present the security result of the architecture against malicious transaction injection into the blockchain. Here, we added a peer (malicious node) that was not part of the blockchain to network I. We assume that an attacker may find its way into joining the blockchain. The malicious node prepares a request transaction and submits it to the blockchain network for verification. Although the transaction agrees with the standard, we observed that a failed transaction notification was sent to the owner. The transaction failed because the sender is not authorized to submit a transaction; hence, it fails the verification step. We manually created the transaction address and used it to check if the transaction is validated and stored in the ledger. The result shows that no transaction is retrieved; hence the transaction is not uploaded to the blockchain network.

2) Insider Threat Detection

Here, we tested the security of the architecture against two typical ways a malicious insider can have unauthorized assess to patient's health records.

a) Multiple Requests: We performed an attack where a malicious insider compromises a peer and begins to send a large amount of what appears to be a legitimate standard formatted request to mount a DoS attack on the blockchain network. The peer prepares a request transaction and submits it to the blockchain. The peer persistently submits the same transactions to mount a DoS attack on the blockchain network. Although other authorized nodes are working to validate the transaction, we observed that the transactions are not validated. The frequency of receiving the same or similar transaction from the same peer exceeds the threshold set in the smart contract. As a result of that, the transactions failed the verification step. We persistently submit the same request from the same authorized node, and we observed that the peers stop mining after the sender was flagged to be compromised. The smart contract automatically drops all subsequent transactions from the same peer.

b) *Unauthorized retrieval of patient's record*: We implement a case where a malicious insider attempts to retrieve patient information. It is assumed that an attacker is unlikely to hold an authorized peer in a compromised state for too long due to frequent security checks. Based on this assumption, an attacker makes all efforts to assess the patient's medical record within the shortest time possible. The result showed that no information was returned because the peer not authorized to retrieve the information. In the smart contract, information retrieval privilege is set for each peer (i.e., the node can only retrieve information that it prepares the request). The architecture drops the query because the node has no retrieval privilege for that patient's EHRs, which makes it suspicious to have been compromised.

V. CONCLUSION

In this on-going research work, we propose a blockchain-based architecture that facilitates and secures inter-healthcare EHRs exchange. In the implementation, each healthcare system maintains patient health information on private blockchain networks. The work focuses on improving the security of both stored and in transit EHRs to strengthen the interoperability between healthcare systems. We present the preliminary result of the security analysis on the detection and prevention of malicious transactions within a healthcare system. The result shows that our architecture has a prospect of detecting and preventing malicious activities from either insider or outsider threats.
As part of the continuation of the work, we wish to expand our work to accommodate the following:

1. Test for more malicious activities within a healthcare system.
2. Evaluate how to detect malicious replies or requests coming for another healthcare system.
3. Evaluate the scalability and its Response time.

REFERENCES

1 J. Clement (2020), Number of U.S. data breaches 2013-2019,

2 Heather Landi (2020), number of patients records breached nearly triples in 2019.

3 V. Liu, MA. Musen, T. Chou, "Data breaches of protected health information in the United States. Jour. of Amer. Med. Assoc. vol 313, num 14, (2015) pp. 1471–1473.

4 Jannetti MC. Safeguarding patient information in electronic health records. AORN vol. 100, num 3, (2014) pp. C7–C8.

5 Hunter, E.S., Electronic health Records in an Occupational Health Setting--Part I. A global overview. *Workplace health & safety*. Vol 61, num 2, (2013) pp.57–60.

6 Lemke J. Storage and security of personal health information. OOHNA J. vol 32, num 1, (2013) pp 25–26.

7 Liu V, Musen MA, Chou T. Data breaches of protected health information in the United States. JAMA. Vol. 313, num 14, (2015) pp 1471–1473.

8 Wang CJ, Huang DJ. The HIPAA conundrum in the era of mobile health and communications. JAMA. Vol. 310, num 11, (2013) Pp.1121–1122.

9 Sultan Aldossary, William Allen. Data Security, Privacy, Availability, and Integrity in Cloud Computing: Issues and Current Solutions. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 4, 2016 pp.485-498

10 C. Wang, S. Chow, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for secure cloud storage," Computers, IEEE Transactions on, vol. 62, no. 2, (2013) pp. 362–375, Feb 2013

11 X. Yang, T. Li, R. Liu, and M. Wang, "Blockchain-Based Secure and Searchable EHR Sharing Scheme," 2019 4th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), Hohhot, China, 2019, pp. 822-8223,

12 X. Zheng, R. R. Mukkamala, R. Vatrapu, and J. Ordieres-Mere, "Blockchain-based Personal Health Data Sharing System Using Cloud Storage," 2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom), Ostrava, 2018, pp. 1-6.

13 S. Amofa et al., "A Blockchain-based Architecture Framework for Secure Sharing of Personal Health Data," 2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom), Ostrava, 2018, pp. 1-6.

14 K. Ito, K. Tago, and Q. Jin, "i-Blockchain: A Blockchain-Empowered Individual-Centric Framework for Privacy-Preserved Use of Personal Health Data," 2018 9th International Conference on Information Technology in Medicine and Education (ITME), Hangzhou, 2018, pp. 829-833.

15 G. Yang and C. Li, "A Design of Blockchain-Based Architecture for the Security of Electronic Health Record (EHR) Systems," 2018 IEEE International Conference on Cloud Computing Technology and Science (CloudCom), Nicosia, 2018, pp. 261-265.

16 P. Zhang, M. A. Walker, J. White, D. C. Schmidt and G. Lenz, "Metrics for assessing blockchain-based healthcare decentralized apps," *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom),* Dalian, 2017, pp. 1-4, doi: 10.1109/HealthCom.2017.8210842

17 X. Liang, J. Zhao, S. Shetty, J. Liu, and D. Li, "Integrating blockchain for data sharing and collaboration in mobile healthcare applications," 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Montreal, QC, 2017, pp. 1-5, doi: 10.1109/PIMRC.2017.8292361.

18 S. Nakamoto (2008) Bitcoin: a peer-to-peer electronic cash system.

19 Zikratov, I., Kuzmin, A., Akimenko, V., Niculichev, V., Yalansky, L.: Ensuring data integrity using Blockchain technology. In: Proceeding of the 20th Conference of fruct Association ISSN 2305-7254 IEEE (2017)

20 M Signorini and M Pontecorvi, W Kanoun, and R Di Pietro, "BAD: a Blockchain Anomaly Detection solution" arXiv:1807.03833v2, [cs. C.R.] 12 Jul 2018

21 T. Golomb, Y. Mirsky, and Y. Elovici " CIoTA: Collaborative IoT Anomaly Detection via Blockchain" arXiv:1803.03807v2, [cs.CY] 09 Apr 2018

22 Gu, J, B Sun, X Du, J Wang, Y Zhuang and Z Wang (2018). Consortium blockchain-based malware detection in mobile devices. IEEE Access, 6, 12118–12128

23 Abdullah, N., Hakansson, A., & Moradian, E. (2017). Blockchain-based approach to enhance big data authentication in distributed environment. In Ubiquitous and future networks (icufn), 2017 ninth international conference on (pp. 887–892).

24 V. Mahore, P. Aggarwal, N. Andola, Raghav and S. Venkatesan, "Secure and Privacy Focused Electronic Health Record Management System using Permissioned Blockchain," 2019 IEEE Conference on Information and Communication Technology, Allahabad, India, 2019, pp. 1-6, doi: 10.1109/CICT48419.2019.9066204.

25 K. Ito, K. Tago and Q. Jin, "i-Blockchain: A Blockchain-Empowered Individual-Centric Framework for Privacy-Preserved Use of Personal Health Data," *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, Hangzhou, 2018, pp. 829-833, doi: 10.1109/ITME.2018.00186.

26 X. Zheng, R. R. Mukkamala, R. Vatrapu and J. Ordieres-Mere, "Blockchain-based Personal Health Data Sharing System Using Cloud Storage," *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)*, Ostrava, 2018, pp. 1-6, doi: 10.1109/HealthCom.2018.8531125.

27 A. Azaria, A. Ekblaw, T. Vieira and A. Lippman, "MedRec: Using Blockchain for Medical Data Access and Permission Management," *2016 2nd International Conference on Open and Big Data (OBD)*, Vienna, 2016, pp. 25-30, doi: 10.1109/OBD.2016.11.

28 G. Carter, H. Shahriar and S. Sneha, "Blockchain-Based Interoperable Electronic Health Record Sharing Framework," 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC), Milwaukee, WI, USA, 2019, pp. 452-457, doi: 10.1109/COMPSAC.2019.10248.

29 O. Ajayi, M. Cherian and T. Saadawi, "Secured Cyber-Attack Signatures Distribution using Blockchain Technology," 2019 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), New York, NY, USA, 2019, pp. 482-488, doi: 10.1109/CSE/EUC.2019.00095.

30 O. Ajayi, O. Igbe and T. Saadawi, "Consortium Blockchain-Based Architecture for Cyber-attack Signatures and Features Distribution," 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York City, NY, USA, 2019, pp. 0541-0549, doi: 10.1109/UEMCON47517.2019.8993036.

Networked Mixed Reality Framework for the Internet of Things

Daniel Wright
Electrical & Computer Eng. (ECE) Dept.
California State University, Fresno
California, USA
wrightdj@mail.fresnostate.edu

Russell Skaggs-Schellenberg
ECE Dept.
California State University, Fresno
California, USA
schellenberg_107066628@mail.fresnostate.edu

Shahab Tayeb
ECE Dept.
California State University, Fresno
California, USA
tayeb@csufresno.edu

***Abstract*—The Internet of Things (IoT) world is developing rapidly, yet seamless integration into our daily lives is lacking. Vendor-specific products tend to restrict customers to their software applications, due to their inherent steep learning curve. We develop the Mixed Reality Control Panel (MRCP) system to address this gap by offering a cross-platform, open-source, vendor-neutral, mixed reality software application capable of interacting with headless, IoT devices. The MRCP application, Smart Parking Meter, and Smart Lock are developed and tested against existing solutions. The duration to perform each action is compared as well as the benefits of additional smart features. Having a mixed reality interface allows for a agnostic and quick interaction with the IoT devices. Early development decisions on cross-platform and vendor-neutral enable future scalability. As a result, both devices are deemed viable solutions for future smart city integrations. Further development will take place for the MRCP system. While others can benefit from the project by incorporating mixed reality as a solution for Smart environments to bridge the gap between IoT and our daily lives.**

***Index Terms*—Augmented Reality, Internet of Things, Mixed Reality, Smart Home, Smart Lock, Smart Parking**

I. Introduction

The proposed MRCP system consists of two main components. The first one being the software application on the user's phone or tablet, which will interact with the second component being the headless devices. In this project two headless devices were developed, a Smart Parking Meter and a Smart Lock. A diagram showing how the smart devices communicate within a network and the application is shown in Fig. 1. The MRCP application utilizes the user's camera to select which smart device to connect to via Bluetooth Low Energy (BLE). The user's device is capable of functioning as a gateway to the web server, or the smart devices can connect to the webserver through a WiFi access point.

II. Background

A. Smart Parking

The application of the MRCP Smart Parking Meter would exist within a Smart City as a Smart Parking solution to a growing problem. Models indicate that 70% of the population will live in urban areas by 2050. Currently such densely

978-1-7281-9615-2/20$31.00 ©2020 IEEE

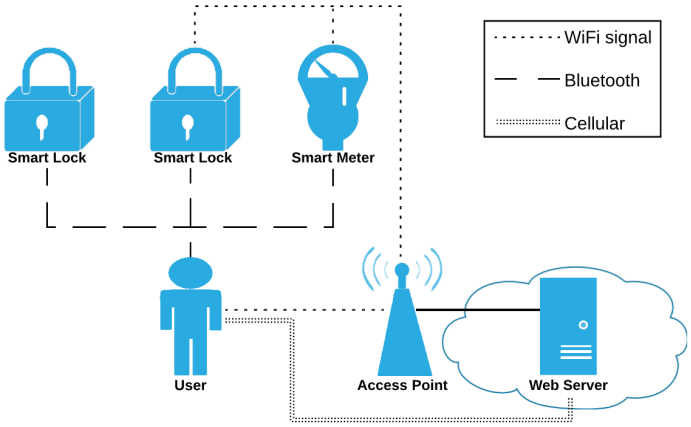


Fig. 1. Communication diagram of the MRCP system.

populated cities consume 75% of the world's energy, while generating 80% of the greenhouse gases. Smart City is a solution that will alleviate some of the impact predicted by population models. A Smart City uses information, digital, and telecommunication technologies with traditional networks and services to improve a city's operations to be more flexible, efficient, and sustainable [1] [2].

A Smart City consists of multiple components, including smart infrastructure, smart buildings, smart transportation, smart energy, smart healthcare, smart technology, smart governance, smart education, and smart citizens [2]. The focus of this section will be within the smart transportation realm, particularly smart parking. Smart parking uses modern technology to identify and relay information about empty/occupied parking spots; This information is then used to create a real-time parking map. This map serves multiple purposes; It can be used to assist drivers with parking, or even report traffic violations [3]. As a result of smart parking, economic, and social benefits emerge. One of these is fuel consumption, As it decreases, so will the cost and impact of transportation. With less congestion from those who are searching for parking spots, the amount of time one spends traveling to their destination is shorter [4]. The main argument against smart parking revolves around privacy concerns. However, with the ubiquitous placement of traffic cameras, this point is somewhat

moot. Therefore we seek to understand how the internet of things can aid with rampant parking problems.

B. Mixed Reality

The MRCP application's interface utilizes Mixed Reality (MR) as an immersive and interactive experience by connecting physical objects to virtual content in a meaningful way [5]. Augmented Reality(AR) is a subcategory of MR. AR is different from Virtual Reality (VR), because it overlays virtual information onto the real world [6]. Comparatively, VR immerses the user into a digital world viewing digital objects. The MRCP application will allow AR interfaces to interact with physical devices. Others have implemented MR and AR into their Smart City solution. Özcan et al combined AR and geolocation to overlay names associated with their campus's buildings [7]. Demir and Karaarslan used their geolocation based AR application GökovAR to promote tourism [8]. Meegahapola and Perera developed the mixed reality application SPMRA to enhance customers in-store shopping experience [9]. Hung et al also created a system using AR to control a door's lock [10].

C. Markers

Marker barcodes were adhered to each device to uniquely distinguish them for communication. An example of a marker used in this project is shown in Fig. 2. These markers consisted of a 4x4 grid allowing 8,192 unique identifiers to be generated. Another option utilizes the 5x5 markers into 4,194,304 unique identifiers [11]. With this in mind, it was determined that digital markers were adequate for this project and immediate future scaling.

Fig. 2. MRCP App marker barcode.

D. Microcontroller

The ESP-32S microcontroller was used to add smart functionality to the electronic lock and for the development of the MRCP Smart Meter. This allowed for additional functionality such as WiFi (IEEE 802.11 b [12]/g [13]/n [14]) and BLE (Bluetooth Special Interest Group 4.0 [15]). There are also 38 available pins for additional components, such as sensors.

III. Design and Development

A. MRCP Application

To make the application cross-platform easily integrate with third-party libraries and services, a web stack was used. This allowed the MRCP application to run in a web browser. Debugging software more efficient because it could be run either on phone or desktop. The majority of the application was written with HTML5, CSS3, and Javascript.

The login screen is shown in Fig. 3 is connected to Firebase and starts a session with the user once they sign in. The browser window is redirected to the main MRCP application page if the email and password for the user are correct.

Fig. 3. MRCP app login page.

The main MRCP application page uses marker barcodes to position the mixed reality interfaces. AR.js is used to convert real-world marker locations to virtual 3-D coordinates and to render the camera output. Once the marker is recognized, Three.js is used to translate the 3-D coordinates to 2-D screen-space coordinates. An HTML interface based on the marker barcode is rendered on the screen at the position of the marker as shown in Fig. 4. The interface tracks the marker in real-time and scales in size depending on the distance of the camera to the marker.

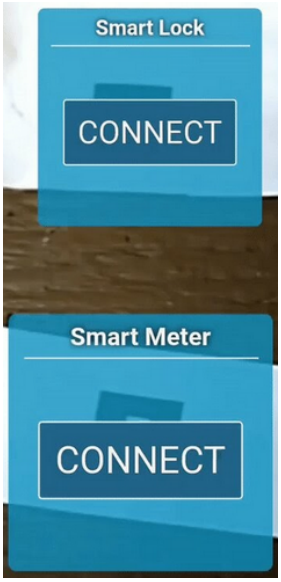


Fig. 4. MRCP app interfaces.

To connect to the MRCP device using BLE, the app utilizes the web Bluetooth that is built into Chrome. When the "CONNECT" button is pushed, the user is asked to pair with a device and the communication protocol begins. Each device type has its interface and communication process that is defined in its Javascript class.

B. MRCP Smart Lock

The Smart Lock's system resides within two states. State 0 (S0) is the state in which the door is locked. It will remain locked until it is manually unlocked or it is digitally unlocked using the MRCP application. S1 is the state in which the door is unlocked. It will remain unlocked unless manually or digitally locked, as well as if the auto-lock timer has expired. This added feature of having an auto-lock can be set to automatically lock the door after a set period. The state diagram of the Smart Lock is shown in Fig. 5.

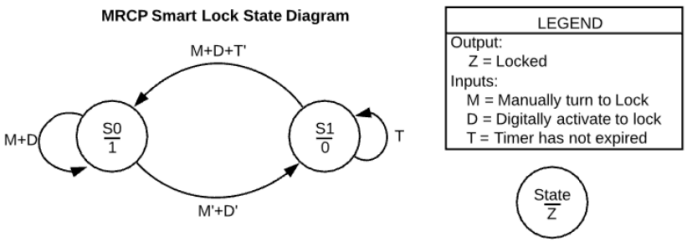


Fig. 5. State diagram of the MRCP Smart Lock.

To implement this system, an electronic lock was purchased shown in Fig. 6 with its unique marker. This lock was originally controlled through a key fob, using Radio Frequency (RF) communication.

Fig. 6. MRCP Smart Lock Exterior Side (Left) and Interior Side (Right).

Since the device needed to communicate via Bluetooth and not RF, the lock was stripped of its electronics and was left with its housing and mechanical functionality. The remaining components consisted of a DC motor, relay switch, gears to drive the deadbolt and housing. The circuitry was then made into a prototype and tested before being soldered on a development board and placed in the housing.

C. MRCP Smart Parking Meter

The functionality of the Smart Parking Meter was determined by first developing the different states the meter would cycle through. These states are displayed in Fig. 7 with their outputs being illuminated with an LED. When no car is detected within a parking spot, it resides in State 0 (S0) as "Available". As soon as a car is detected, it transitions into State 1 to "StandBy". It waits for a designated amount of time to allow the user to connect to the smart meter via the MRCP app. After the user connects and checks in, the smart meter transitions into State 3 as "Valid Parking" and keeps track of the amount of time the user is parked there. Once the meter detects that the vehicle no longer resides within the spot using its ultrasonic range sensor, it ends the session with the user then transitions back into the "Available" state. If the user does not check-in with the meter before the grace period has expired, the meter transitions into State 4 to report a "Parking Violation". At this time the LED rapidly blinks red, signaling that a parking violation is taking place. Lastly, State 2 and State 5 represent "Maintenance Required Valid Parking" and "Maintenance Required Invalid Parking" respectively. These states are used when the sensor is blocked or requires maintenance to function properly.

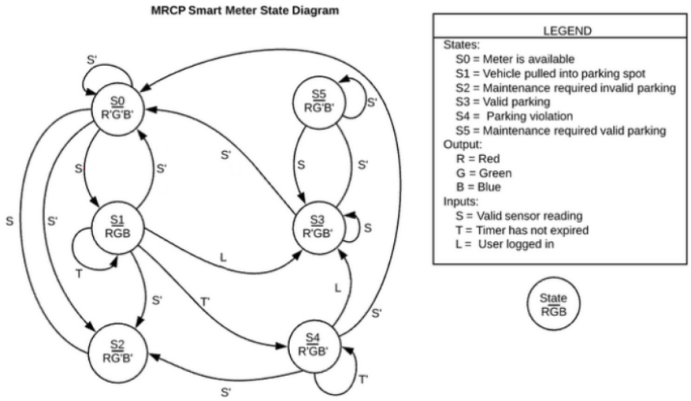


Fig. 7. MRCP Smart Parking Meter state diagram.

D. Smart Parking Network

For the MRCP Smart Meter, a network suitable needed to be developed and analyzed. Researching a variety of models led to Sadhukhan's proposed IoT approach to creating an E-Parking system [16]. This system implemented smart Parking Meters (PM) which contained an ultrasonic sensor to detect whether a parking spot was reserved or available. This information was transmitted via WiFi (IEEE 802.11 b/g/n) to Access Points (AP) throughout the parking lot. The AP would relay the information to a Local Parking Management System (LPMS) which functioned as an internet gateway to send data to a web.

Sadhukhan's proposed E-Parking system was a good fit for the MRCP Smart Parking Meter network. In particular, the method in which PM communicates with the AP and the AP communicates with a LPMS. This section would utilize the ESP-32S, the same affordable, low energy microcontroller capable of communicating via WiFi (IEEE 802.11 b/g/n) and Bluetooth (SIG 4.0) used in the MRCP Smart Parking Meter. These motes would transmit sensor readings through a centralized network forwarding their data to a Cisco 1941W

Integrated Services Router. The router, in this case, would function as an AP, sending the received data to a LPMS. By developing this link, new PMs will be capable of connecting to their system's web server.

The ESP-32S utilizes its WiFi module, which uses IEEE 802.11 b/g/n standards. This protocol was used in Sadhukhan's model as well as various other models [17], [18]. Although Bluetooth offered an appealing low energy option [19], it was dismissed for the networking aspect because of the limited connections allowed [20], [21] as requiring additional hardware to construct a gateway.

The system is capable of scaling to accommodate the size of any parking lot. This can be achieved by adding more wireless routers and configuring their routing tables. To ensure that the designed system is functioning properly, multiple Quality of Service (QoS) metrics would need to be tested. Singh and Baranwal discuss the different QoS metrics that should be considered when analyzing an IoT system [22]. The metrics considered when testing the system were network focused. Response times were tested, to observe how long a mote requires to detect a change in a parking spot, and to reflect the change in a web server. The capacity of each access point could be tested to ascertain a limit of how many motes an AP and its network would be able to reliably support. Though this would be an expensive experiment to perform till failure, the affects of adding motes could be analyzed.

IV. The Proposed System

A. MRCP Smart Lock

The Smart Lock functioned as intended, with the capability to lock and unlock the door through the MRCP application. At the same time, it maintained its functionality to manually lock and unlock. Fig. 8 consists of a set of images that chronologically showcases the MRCP Lock unloading when the user pressed the "UNLOCK" button on the application.

Fig. 8. MRCP Lock transitioning from State 0 to State 1.

B. MRCP Smart Parking Meter

The MRCP Smart Parking Meter's housing was modeled, then 3D printed. The ultrasonic sensor and LED were then installed, followed by the circuitry and marker. Black iron pipe and flange was used to elevate and support the meter. The final form of the smart meter is shown in Fig. 9.

Fig. 9. Completed MRCP Smart Parking Meter.

A chronological set of images demonstrate the functionality of the interface in Fig. 10- 12. The first figure displays the interface with a button labeled "Connect" which when pressed connects the user to that MRCP Smart Parking Meter.

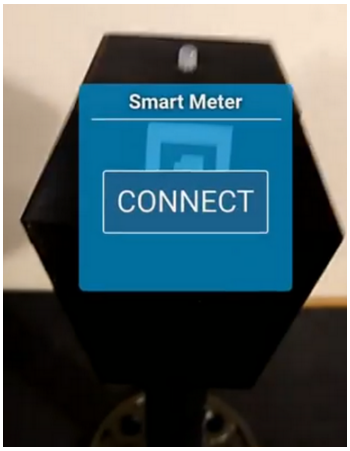


Fig. 10. MRCP Smart Parking Meter connection interface.

Once connected, the interface transitions to notify the user of the rate to park in the current spot as shown in 11. In this case, the rate is configured at $5 per hour. The user can accept these terms by selecting the "Start" button and begin a valid parking session or they can vacate the parking spot. Doing neither of these options will result in a parking violation, in which the meter will transition into State 4.

After accepting the terms, the interface transitions into its final stage displaying the time duration the vehicle has been parked, as well as the accumulated cost to the user. When the meter detects that the vehicle has left the parking spot, the session with the user ends and shifts back to State 0.

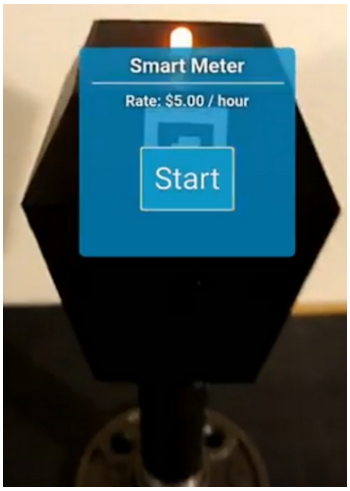


Fig. 11. MRCP Smart Parking Meter accept rate interface.

Fig. 12. MRCP Smart Parking Meter current duration and total interface.

C. Smart Parking Metrics

The ESP-32S's hall effect sensor was configured to reflect a sensor that would detect the presence of a vehicle. When triggered, the availability of that spot is no longer available. Every few seconds, the sensor is sampled, and an HTTP POST is transmitted through the AP to a server. The server was originally hosted on a website, however it became apparent during testing that the desired packets would not be accessible in Wireshark. The server was then hosted on a local machine on the same network which allowed the packets to be observed.

The first test observed how much time it took the mote to update the database. To determine this, the packet size was obtained while monitoring traffic through Wireshark. It was noted that the length of the packet was 95 bytes. To calculate the average time the packet would take to reach the server, ICMPv4 was used. Fig. 13 displays the output from the ping, in which 20 packets were sent with their size matching the original 95 bytes.

Additional motes were added to the system and the test was

```
user@UsersPC:~$ ping -c 20 -s 95 -i 2 192.168.1.4
PING 192.168.1.4 (192.168.1.4) 95(123) bytes of data.
103 bytes from 192.168.1.4: icmp_seq=1 ttl=255 time=102 ms
103 bytes from 192.168.1.4: icmp_seq=2 ttl=255 time=4.57 ms
103 bytes from 192.168.1.4: icmp_seq=3 ttl=255 time=76.6 ms
103 bytes from 192.168.1.4: icmp_seq=4 ttl=255 time=258 ms
103 bytes from 192.168.1.4: icmp_seq=5 ttl=255 time=40.5 ms
103 bytes from 192.168.1.4: icmp_seq=6 ttl=255 time=217 ms
103 bytes from 192.168.1.4: icmp_seq=7 ttl=255 time=262 ms
103 bytes from 192.168.1.4: icmp_seq=8 ttl=255 time=192 ms
103 bytes from 192.168.1.4: icmp_seq=9 ttl=255 time=11.2 ms
103 bytes from 192.168.1.4: icmp_seq=10 ttl=255 time=64.2 ms
103 bytes from 192.168.1.4: icmp_seq=11 ttl=255 time=209 ms
103 bytes from 192.168.1.4: icmp_seq=12 ttl=255 time=29.5 ms
103 bytes from 192.168.1.4: icmp_seq=13 ttl=255 time=33.6 ms
103 bytes from 192.168.1.4: icmp_seq=14 ttl=255 time=723 ms
103 bytes from 192.168.1.4: icmp_seq=15 ttl=255 time=89.4 ms
103 bytes from 192.168.1.4: icmp_seq=16 ttl=255 time=11.1 ms
103 bytes from 192.168.1.4: icmp_seq=17 ttl=255 time=58.4 ms
103 bytes from 192.168.1.4: icmp_seq=18 ttl=255 time=656 ms
103 bytes from 192.168.1.4: icmp_seq=19 ttl=255 time=27.9 ms
103 bytes from 192.168.1.4: icmp_seq=20 ttl=255 time=9.36 ms

--- 192.168.1.4 ping statistics ---
20 packets transmitted, 20 received, 0% packet loss, time 38033ms
rtt min/avg/max/mdev = 4.579/153.931/723.616/197.581 ms
```

Fig. 13. ICMPv4 test to acquire average time of HTTP packet.

repeated. The amount of time it took a mote to ping the server increased as more motes connected to the server. The result was graphed below in Fig. 14.

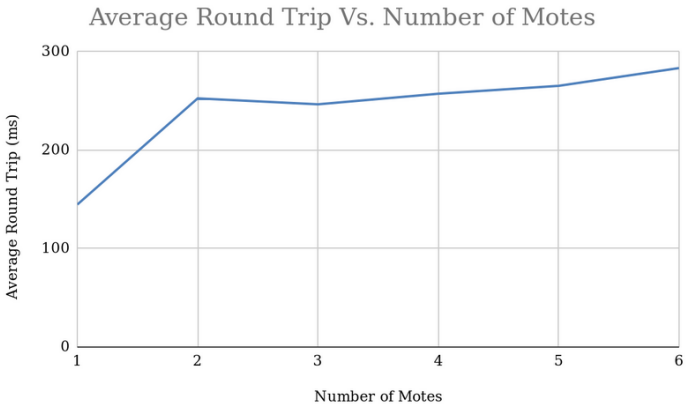


Fig. 14. ICMPv4 test to acquire average time of HTTP packet.

Analyzing the data collected from the results, it was determined that the system was capable of supporting the proposed methods. The system was capable of sensing a change in its magnetic field and relaying information to the server to display that data. The amount of time it took a mote to relay its sensor reading was between 4.579-723.616ms, with the average time at 153.193ms as shown in Figures 4 and 5. This was enough speed to accurately support the system.

The amount of time it took for a mote to perform a TCP handshake was 0.5226741s. This was calculated by subtracting the start time of the handshake by the end time, which was collected in Wireshark. This was an adequate speed to support the system.

The amount of time it took a mote to ping the server increased as more motes were added to the network. Since each mote needed to consistently keep in contact with the server, the number of packets the access point required to process increased.

Lastly, the average round trip to communicate from the server to a mote was analyzed. During this test, 20 packets were sent with their time-averaged. This was performed at 10 ft intervals, down a hallway with cement walls. The data gathered was inconclusive as shown in Fig. 15 among various test runs. It was determined that the environment was most likely the cause of the inconsistent data, with the WiFi signal reflecting off of the walls. Although all ranges were within the capability of the AP's 802.11n [23], [24], its speed was expected to decrease while increasing distance [25]. Regardless, the packets were reliably transmitted within the measured ranges, and it was not until the mote was moved behind a corner that the signal was lost. An experiment testing the distance in an ideal environment like an actual parking lot would more accurately test our system's range.

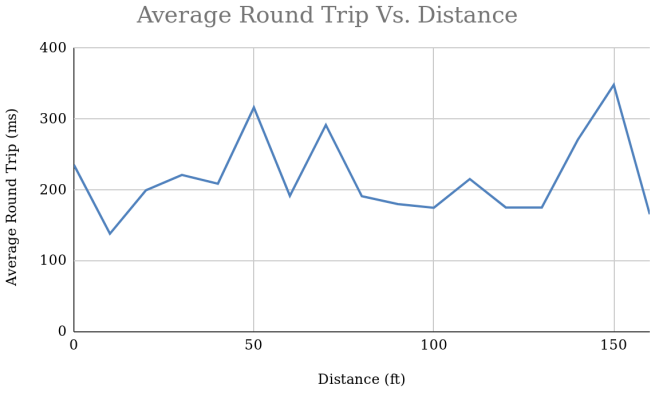


Fig. 15. Graph of round trip time over a range of distances.

Determining the range of the system would be beneficial in concluding how many parking spots each AP would support. According to the data results, that the mote could communicate effectively up to 160 ft. Implemented inside a parking lot, Fig. 16 illustrates the coverage that would be expected.

Ideally the AP would be able to cover a wider range, considering the approximate range of 802.11n is 400ft for outdoor usage [24]. With the increase of coverage, the network's subnet would need to change to accommodate the extra motes.

The amount of time to activate an MRCP Smart Parking Meter against its competitors was also analyzed. Table I below displays the meters being evaluated and the amount of time it takes to reserve a parking spot for one hour and the payment method used. The time required to use a meter was determined by analyzing videos where similar meters were used.

It was determined that interfacing through the MRCP system was the most efficient when reserving a parking space. It took an average of 6 seconds to reserve a space, where most of that time was spent connecting to the device. Future development with the application will reduce this connection further. Using the IPS system with a card payment required more interfacing to get the appropriate amount of time which took a total of 12 seconds. While using the same system with manual coin payment, this required inputting eight quarters

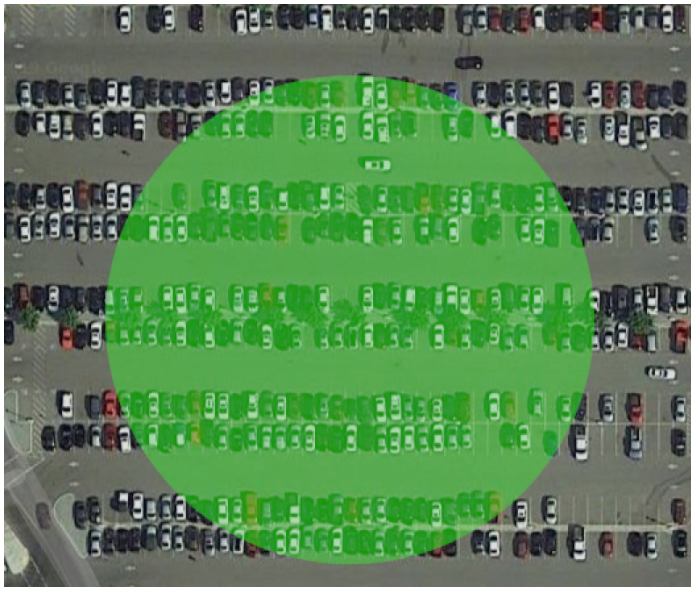


Fig. 16. Coverage area an AP can expect with the tested range of 160ft.

TABLE I
COMPARISON OF SMART METERS ACTION TIME.

Meter	Payment	Action Time (Seconds)
MRCP	Application	6
IPS	Card	12
IPS	Coins	20
GeoPark	Card	32
PayByPhone	Application	12

which took 20 seconds [26]. GeoPark and PayByPhone were also studied, which are multi-space meters. GeoPark requires the user to enter their parking space number and duration of stay. The meter would then print out a permit for the user to stick in their windshield. This transaction took a total of 32 seconds at the shared meter, not accounting for the commute necessary to access the meter [27]. The other option PayByPhone offered a similar option to dispense a permit, but also had the option to use their application. Using their application required that the parking meter number be entered, but stored the user's vehicle information to save time. However, this transaction still took 12 seconds to complete [28].

The time necessary to interact with each meter should not be the only variable considered. While PayByPhone did prompt the user to extend their session via their application whenever their allotted time was about to expire, the MRCP offered a better solution by only charging the user for the duration of their stay, and not requiring additional confirmations. A beneficial feature that the MRCP and PayByPhone system shared was their interactions did not require any physical contact with publicly shared objects.

V. Summary

The MRCP system provides a viable solution for users to interact with physical IoT devices. By introducing a mixed reality interface, users will naturally interact with the IoT world, as they intended. A Smart Parking Meter and Smart Lock were developed for use under the MRCP application as a proof of concept.

Acknowledgment

We would like to thank Dr. Nan Wang for his advice in this project, Roberto Reyes for his assistance in the MRCP's networking section. This project was partially funded by a scholarship funded by DPS Telecom.

References

[1] S. Tayeb, S. Latifi, and Y. Kim. "A survey on IoT communication and computation frameworks: An industrial perspective." 2017 IEEE 7th annual Computing and Communication Workshop and Conference (CCWC). IEEE, 2017. doi: 10.1109/MCE.2016.2556879"

[2] Tayeb, S., Raste, N., Pirouz, M., and Latifi, S. (2018). A Cognitive Framework to Secure Smart Cities. In MATEC Web of Conferences (Vol. 208, p. 1).

[3] "What is Smart Parking, and Who Benefits?," RSS. [Online]. Available: https://www.leverege.com/blogpost/what-is-smart-parking-and-who-benefits. [Accessed: 17-Oct-2019].

[4] "R. Mangiaracina, A. Tumino, G. Miragliotta, G. Salvadori and A. Perego, ""Smart parking management in a smart city: Costs and benefits,"" 2017 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), Bari, 2017, pp. 27-32. doi: 10.1109/SOLI.2017.8120964"

[5] K. Cheng and I. Furusawa, "The Deployment of a Mixed Reality Experience for a Small-Scale Exhibition in the Wild," 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Munich, Germany, 2018, pp. 214-215. doi: 10.1109/ISMAR-Adjunct.2018.00070

[6] Feng Zhou, H. B. Duh and M. Billinghurst, "Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR," 2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, Cambridge, 2008, pp. 193-202. doi: 10.1109/ISMAR.2008.4637362

[7] U. Özcan, A. Arslan, M. İlkyaz and E. Karaarslan, "An augmented reality application for smart campus urbanization: MSKU campus prototype," 2017 5th International Istanbul Smart Grid and Cities Congress and Fair (ICSG), Istanbul, 2017, pp. 100-104, doi: 10.1109/SGCF.2017.7947610.

[8] Ö. F. Demir and E. Karaarslan, "Augmented reality application for smart tourism: GökovAR," 2018 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG), Istanbul, 2018, pp. 164-167, doi: 10.1109/SGCF.2018.8408965.

[9] L. Meegahapola and I. Perera, "Enhanced in-store shopping experience through smart phone based mixed reality application," 2017 Seventeenth International Conference on Advances in ICT for Emerging Regions (ICTer), Colombo, 2017, pp. 1-8, doi: 10.1109/ICTER.2017.8257810.

[10] C. Hung, Y. Fanjiang, K. Chung and C. Kao, "A door lock system with augmented reality technology," 2017 IEEE 6th Global Conference on Consumer Electronics (GCCE), Nagoya, 2017, pp. 1-2, doi: 10.1109/GCCE.2017.8229462.

[11] Marker generator. [Online]. Available: https://au.gmented.com/app/marker/marker.php. [Accessed: 19-Mar-2020].

[12] IEEE 802.11b-1999 - IEEE Standard for Information Technology - Telecommunications and Information Exchange between Systems - Local and Metropolitan Networks - Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher Speed Physical Layer (PHY) Extension in the 2.4 GHz Band. C/LM - LAN/MAN Standards Committee, June 2003, URL: https://standards.ieee.org/standard/802_11b-1999.html.

[13] IEEE 802.11g-2003 - IEEE Standard for Information Technology– Local and Metropolitan Area Networks– Specific Requirements– Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Further Higher Data Rate Extension in the 2.4 GHz Band. C/LM - LAN/MAN Standards Committee, 2003, URL: https://standards.ieee.org/standard/802_11g-2003.html.

[14] IEEE 802.11n-2009 - IEEE Standard for Information Technology– Local and Metropolitan Area Networks– Specific Requirements– Part 11: Wireless LAN Medium Access Control (MAC)and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput. C/LM - LAN/MAN Standards Committee, 29 Oct. 2009, URL: https://standards.ieee.org/standard/802_11n-2009.html.

[15] Bluetooth SIG (2010) Specification of the Bluetooth-System - Covered Core Package version: 4.0 URL: https://www.bluetooth.com/specifications/archived-specifications/

[16] P. Sadhukhan, "An IoT-based E-parking system for smart cities," 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi, 2017, pp. 1062-1066. doi: 10.1109/ICACCI.2017.8125982

[17] C. Yuan, L. Fei, C. Jianxin and J. Wei, "A smart parking system using WiFi and wireless sensor network," 2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), Nantou, 2016, pp. 1-2. doi: 10.1109/ICCE-TW.2016.7520924

[18] K. Hassoune, W. Dachry, F. Moutaouakkil and H. Medromi, "Smart parking systems: A survey," 2016 11th International Conference on Intelligent Systems: Theories and Applications (SITA), Mohammedia, 2016, pp. 1-6. doi: 10.1109/SITA.2016.7772297

[19] G. D. Putra, A. R. Pratama, A. Lazovik and M. Aiello, "Comparison of energy consumption in Wi-Fi and bluetooth communication in a Smart Building," 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, 2017, pp. 1-6. doi: 10.1109/CCWC.2017.7868425

[20] Kettimuthu, Rajkumar and Sankara Muthukrishnan. "Is Bluetooth suitable for large-scale sensor networks?" ICWN (2005).

[21] J. Lee, Y. Su and C. Shen, "A Comparative Study of Wireless Protocols: Bluetooth, UWB, ZigBee, and Wi-Fi," IECON 2007 - 33rd Annual Conference of the IEEE Industrial Electronics Society, Taipei, 2007, pp. 46-51. doi: 10.1109/IECON.2007.4460126

[22] M. Singh and G. Baranwal, "Quality of Service (QoS) in Internet of Things," 2018 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU), Bhimtal, 2018, pp. 1-6. doi: 10.1109/IoT-SIU.2018.8519862

[23] "Cisco 1941 Series Integrated Services Routers Data sheet," Cisco, 22-Aug-2017. [Online]. Available:https://www.cisco.com/c/en/us/products/collateral/routers/1900-series-integrated-services-routers-isr/data_sheet_c78_556319.html. [Accessed: 14-Dec-2019].

[24] N. Tengyuen, "12 Wireless Router Antenna Distance Coverage Comparison," GeckoandFly, 08-Nov-2019. [Online]. Available: https://www.geckoandfly.com/10213/wireless-router-antenna-distance-coverage-comparison/. [Accessed: 14-Dec-2019].

[25] A. Matsumoto, K. Yoshimura, S. Aust, T. Ito and Y. Kondo, "Performance evaluation of IEEE 802.11n devices for vehicular networks," 2009 IEEE 34th Conference on Local Computer Networks, Zurich, 2009, pp. 669-670. doi: 10.1109/LCN.2009.5355054

[26] "How to Use the Smart Parking Meters in Phoenix - YouTube". [Online]. Available: https://www.youtube.com/watch?v=7j4TMzuF08A. [Accessed: 25-Jul.-2020].

[27] "How to use GeoPark , smart Parking Meters. - YouTube". [Online]. Available: https://www.youtube.com/watch?v=bgj9PM_073E. [Accessed: 25-Jul.-2020].

[28] "Pay by phone demo - YouTube". [Online]. Available: https://www.youtube.com/watch?v=Reg9bBas3ag. [Accessed: 25-Jul.-2020].

PoRCH: A Novel Consensus Mechanism for Blockchain-Enabled Future SCADA Systems in Smart Grids and Industry 4.0

Md Tamjid Hossain, Shahriar Badsha, Haoting Shen
University of Nevada, Reno, NV, USA
Email: mdtamjidh@nevada.unr.edu,{sbadsha, hshen}@unr.edu

***Abstract*—Smart Grids and Industry 4.0 (I4.0) are neither a dream nor a near-future thing anymore, rather it is happening now. The integration of more and more embedded systems and IoT devices is pushing smart grids and I4.0 forward at a breakneck speed. To cope up with this, the modification of age-old SCADA (Supervisory Control and Data Acquisition) systems in terms of decentralization, near-real-time operation, security, and privacy is necessary. In this context, blockchain technology has the potential of providing not only these essential features of the data acquisition process of future SCADA systems but also many other useful add-ons. On the other side, it is evident that various type of security breach tends to take place more during any economic turmoil. These can cause even more serious devastation to the global economy and human life. Thus, it is necessary to make our industries robust, automated, and resilient with secured and immutable data acquiring systems. This paper deals with the implementation scopes of blockchain in the data acquisition part of SCADA systems in the area of the smart grid and I4.0. There are several consensus mechanisms to support blockchain integration in the field cryptocurrencies, vehicular networks, healthcare systems, e-commerce, etc. But little attention has been paid for developing efficient and easy to implement consensus mechanisms in the field of blockchain-enabled SCADA systems. From this perspective, a novel consensus mechanism, which we call PoRCH (Proof of Random Count in Hashes), with a customized mining node selection scheme has been proposed in this paper. Also, a small-scale prototype of a blockchain-enabled data acquisition system has been developed. The performance evaluation of the implemented prototype shows the benefits of blockchain technology.**

***Index Terms*—Smart Grid, Industry 4.0, IoT, SCADA, Security, Blockchain, PoRCH**

I. Introduction

From 2011, when it first appeared in a high-tech strategy project of the German government, till today Industry 4.0 (simply I4.0 or I4.0) has crossed a long path to become a realization from only a concept. To fulfill the objective of creating self-optimized, self-cognitive, and self-customized industries, I4.0 has been incorporated with CPS (Cyber-Physical Systems), IoT (Internet of Things), IoS (Internet of Service) and Smart factory. However, one thing which is common and very important to all of these new technologies to be interconnected and updated for the successful operation of I4.0 is the continuous, secure, and automated flow of data. At the same time, the data acquisition process needs to be trustworthy and faster without the presence of a third-party to prevent unauthorized data manipulation. On the other hand, blockchain technology can be described as an open and distributed ledger that records transactions between two parties efficiently and in a verifiable and permanent way [1]. By design, blockchain is immutable and offers non-repudiation. For these certain characteristics and potentiality, nowadays blockchain technology has become very popular to be implemented in the data acquisition part of the future SCADA systems of the smart grid industry. This can be a giant leap towards the fourth industrial revolution. The motivations behind this research are –

- The data acquisition process needs to be secured, automated, quick, error-free, and robust for any SCADA system which is used in smart grids and other industries.
- Blockchain technology can help the data acquisition process to be a smart fit in the context of I4.0
- Traditional consensus mechanisms of blockchain technology cannot serve efficiently for all the data acquisition process. Hence, sometimes, a customized consensus mechanism is necessary for a private and permissioned type of blockchain network
- Earlier researches related to the mining node selection process can be impractical in real-life scenarios. So, a random and fair mining node selection process is necessary

This major focus point of this paper is on a very important part of the blockchain technology - the consensus mechanism, from the perspective of the data acquisition process of future smart grids. The paper also discusses other essential steps of a private blockchain creation including, but not limited to, hashing mechanism, block creation, verification, and addition. Moreover, a customized mining node selection procedure has been proposed, designed, and demonstrated in the context of the SCADA system of a smart grid. The proposed PoRCH mechanism counts a random number appearance in the hash value of the measurement data and selects the node/server as a mining node that has the largest count. In case of not having a single node bearing the largest count of random number appearance, the scheme proposes to use cryptographically secure random node selection algorithm for mining purpose from all available nodes. In this way, the mechanism ensures

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

randomness and fairness in the mining node selection process of a private blockchain network. The performance evaluation shows that the entire process requires very low computational overhead while preserving data security, privacy, and trust. In a nutshell, the major contributions of the paper are –

- The paper discusses the previous research works on the consensus mechanism of the blockchain-enabled data acquisition system and points out the research gaps
- It bridges one of the common and popular ICS (Industrial Control System) protocols, DNP3 (Distributed Network Protocol 3), with the blockchain network
- The paper proposes a novel consensus mechanism called PoRCH and present a simplified demo by including a customized mining node selection procedure for a private and permissioned blockchain-enabled data acquisition system where incentive or penalty is not required for the validators/miners.

In the remainder of this paper, we have used the words – 'nodes", 'servers', 'field devices', 'relay servers' alternately depending upon the context, but they are implemented as one type of entity that can both transmit and receive data.

II. Related works

In [2], blockchain has been applied to the ICS network of device nodes where every device node executes the smart contract and records the transaction data. Any device nodes on the ICS network can record all transaction data. However, it may then possible for an adversary to compromise a device node and manipulate the data. So, there has to be a novel and decentralized way of tackling cyber-attacks before they take place and cause a single-point-of-failure and data manipulation. Taking this into account, researchers of [3] propose to enhance the integrity and confidentiality of synchrophasor communication networks in an ICS environment through blockchain and essential cryptographic tools (e.g. bloom filter, ECDSA, etc.). The researchers of [4] proposes a scheme- BUS where IoT devices send encrypted data to UAV (Unmanned Aerial Vehicle), and UAV validates the sender's identity using a bloom filter. UAV then sends the data to the nearest server who prepares the data to add to the blockchain. Several other validators also validate the block before adding it to the blockchain. In other researches, blockchain has been applied towards data security in ICS [5], [6]. Reference [5] uses ethereum blockchain in its prototype implementation and may suffer from the problems of ethereum blockchain such as scalability and latency.

Reference [6] uses a customized mining node selection procedure that follows the process of selecting a node as a mining node if its energy consumption data is the closest to the average. However, in a practical scenario, the electricity consumption behavior of any user may not change rapidly and remain the same at least for a while. Thus, it may possible for any node to have the almost same data for concurrent time slots and so, there may be a node whose data will be closest to average for these concurrent time slots. As a consequence, a node may remain as a mining node for a considerable amount of time. If an adversary, somehow, finds out the mining node and compromise it, he/she can exploit this time slots to design and execute his/her attack. Reference [7] addresses the problem of reaching consensus in distributed social networks. However, they have not incorporated the mining technique, rather they've used wired connected RSUs (roadside units). The consensus mechanism is lightweight but may not be applied to the blockchain-based smart grid data acquisition system or any system that requires a near-real-time operation through wireless connectivity. A novel consensus mechanism named PoAh (Proof-of-Authentication) has been proposed in [8]. This consensus mechanism introduces the trusted party in the validation of blocks and solves many problems of traditional consensus mechanisms. However, these schemes have not bridged the OT protocols with the communication network of the blockchain which in turn, has motivated the authors to proposed the PoRCH scheme that aims to connect the dots in this domain.

III. Main Design

A. Communication Protocol

The communication system of an ICS can be broadly categorized into IT (Information Technology) and OT (Operational Technology). IT systems use ERP systems that mostly uses SMTP, SNMP, OPC, SMB, HTTP, XML protocols whereas OT systems encompass MES, DCS, SCADA, RTUs, PLCs, sensors, etc. and use MODBUS, PROFIBUS, PROFINET, DNP3, EtherNet/IP, RS-232, etc.

Here, in this paper, DNP3m (DNP3 minus, a simplified version of DNP3) protocol has been used to collect the data by the control center and data aggregator node from relay nodes. It is widely used in the US power grid infrastructure. It is an application layer protocol built on top of the TCP protocol [9]. The Sample request and response packets are also depicted in Fig. 1. However, the goal of this paper is not to

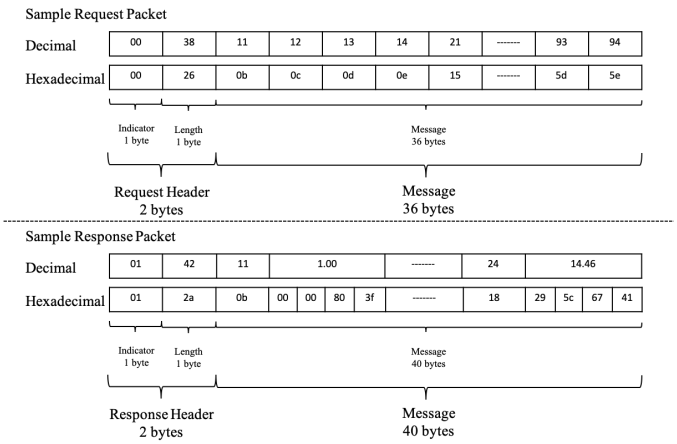


Fig. 1. Sample request and response packet.

find out the most secure communication protocol and discuss in-depth but to implement blockchain technology to make the data acquisition system decentralized, secured, private, and

trustworthy. In the application layer, both the request and response header contains 2 bytes. The first byte represents the identity of the protocol, i.e. the protocol is DNP3m. In this implementation, the first byte is set to '0' for request header and '1' for response header. The second byte reflects the length of the entire message. Then, the message is incorporated with the headers.

B. Relay Servers

Relay servers are the primary sources of measurement data. They are associated with different field devices, sensors, actuators, PMUs, valves. In the prototype demonstration, four nodes/servers have been used to represent the relay servers. The relay servers can act both as servers and clients based on operational needs. Here, relay servers use the DNP3m protocol to send the response with measurement data. UTF-8 (8-bit Unicode Transformation Format) has been used as an encoding mechanism for sending string responses. Another important task of the relay servers is to take part in the voting process of mining node selection. To cast vote for the appropriate node during the mining node selection process, the relay servers first use hashing technique to hash their measurement data. Then, they count the appearance of the random number in their hash values and finally votes for the node to be selected as the mining node who has the largest count.

C. Data Aggregator (DA)

Data Aggregator (DA) plays the most vital role in the proposed data acquisition system. It is responsible for collecting the data, then preprocessing it and coordinate mining node selection and all kinds of the verification process. DA should preferably reside inside the internal network with the relay servers for efficient operation and low computational overhead and time. In this demonstration, DA works as both in server and client mode. As soon as it gets a data acquisition request from CC (Control Center), it starts its operation to collect the data from relay servers and finally completes its operation for a cycle by sending updated blockchain to CC and other relay servers. Another important task of DA is to generate random numbers for assisting the mining node selection procedure. Moreover, it might be the case to have multiple nodes having the largest and same appearance counts of the random number in their hash values or the case to have all nodes having the count as zero (0). In these scenarios, the DA delegates one node from all available nodes based on a random selection algorithm.

D. Control Center (CC)

The CC is responsible for initializing and controlling the operation. CC can be any physical or cloud server and may have HMI (Human Machine Interface) attach with it for better user experiences. At first, CC sends the data acquisition command to DA and finally receives the updated copy of the blockchain containing measured data at the end of each cycle. For simplicity, here in this demonstration, each data acquisition cycle has been considered as 15 seconds. Of course, this period is configurable according to the requirement of the data acquisition system, data generation, and processing time.

E. Blockchain

Blockchain is an open, distributed, and immutable ledger system that was first conceptualized as a public transaction ledger of the cryptocurrency bitcoin by Satoshi Nakamoto in 2008. Since then, blockchain technology has been proposed to apply in several domains, including but not limited to, task allocation [10], ride-sharing [11], cybersecurity information sharing [12], cyber insurance [13] etc. In this paper, we have considered extending the application of blockchain in the smart grid area.

Blockchain implementation starts with the selection of a mining node and ends with the addition of new blocks in a repetitive manner. In this part, data preprocessing, data aggregation, existing blockchain verification, hashing, new block creation (see Fig. 2), block verification, block addition occurs as an intermediate process. Design of the proposed blockchain can be well described by the following processes.

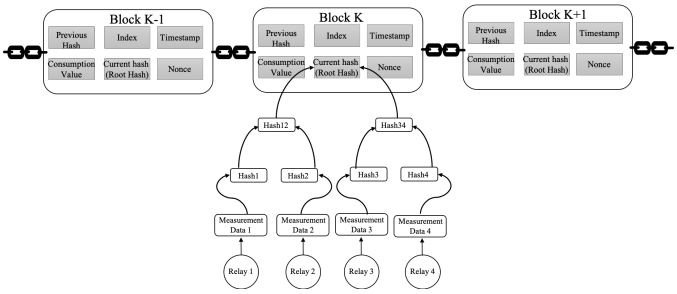


Fig. 2. Example blockchain.

1) Hashing Mechanism: It refers to a process of generating an output of a fixed or predefined length regardless of the length of the input. The result of this computational function is called a hash. In a blockchain, the process of using a specific hash function to process a transaction or data is called hashing.

Rarely, a hash collision may occur as a result of the 'Birthday Box' incident [14]. A hash collision is defined as an instance in which two or more observations are hashed to the same value [15]. In this demonstration, SHA-256 has been implemented as a hashing algorithm by using Python 3.0 'hashlib' module.

2) Consensus Mechanism: At a high level, a consensus mechanism can be described as the process of adding newly published blocks in the blockchain after testing the validity by validators/miners to ensure the trust in the network. Over the years, numerous consensus mechanisms have been proposed, developed, and implemented in several blockchain networks. Both public and private blockchain use consensus mechanism. Popular consensus mechanisms are PoW (Proof of Work), PoS (Proof of Stake), DPoS (Delegated Proof of Stake),

PBFT (Practical Byzantine Fault Tolerance), PoA (Proof of Authority), RAFT [16], etc. Each one has its advantages and drawbacks. PoW requires high computational overheads, PoS is biased to the wealthiest validators and thus unfair for new participants, DPoS is less decentralized and less resilient, PBFT only works on a permissioned blockchain due to the lack of anonymity [17]. Though the authors of [18] have proposed a consensus mechanism named PoD (Proof of Driving) which aims to improve PBFT in VANET (Vehicular Ad hoc Network) application, they have mainly focused on implying it in a public blockchain network. However, in a private blockchain that may be adopted by a small industry and where any incentive or penalty mechanism may not necessary, above mentioned consensus mechanisms would not suit properly due to their drawbacks. So, some sort of customized and lightweight consensus mechanism like the proposed PoRCH mechanism would be a better choice.

Mining node selection: Considering the drawbacks of traditional consensus mechanisms, a customized mining node selection procedure has been designed and demonstrated (see Fig. 3 and algorithm 1). First, the DA server requests data from the relay servers. The relay servers then send the encrypted measurement data to DA. After getting the measurement data from each relay, DA generates a random number and broadcasts this random number to all the relay servers. Relay servers then hash their measurement data and find out that random number appearance count from their respective hash values. Next, the relay servers share the count among each

Algorithm 1: Mining node selection

Input : Measurement Data as D_m, Relay Name as N_r, Random Number from DA as R
Output: The name of the mining node

Function *Generate_Hash(D_m)* **:**
  $hash \leftarrow$ hashlib.sha256(block).hexdigest()
  return $hash$;
end

Function *Count($hash$, R)* **:**
  $valCount \leftarrow hash$.count(R))
  return $valCount$;
end

Set Dictionary: $R_v \leftarrow \{N_r : valCount\}$
Sort Dictionary: $S_{R_v} \leftarrow$ sorted(R_v.items(), key=lambda item: item[1], reverse=True)
number of largest count: $LarCount \leftarrow$ list(S_{R_v}.values()).count(S_{R_v}[list(S_{R_v}.keys())[0]])

if $LarCount == 0$ **then**
  DA selects MiningNode Randomly ;
else if $LarCount == 1$ **then**
  MiningNode $\leftarrow list(S_{R_v})[0]$;
else if $LarCount > 1$ **then**
  DA selects MiningNode Randomly ;
else
  Stop the selection process ;
end

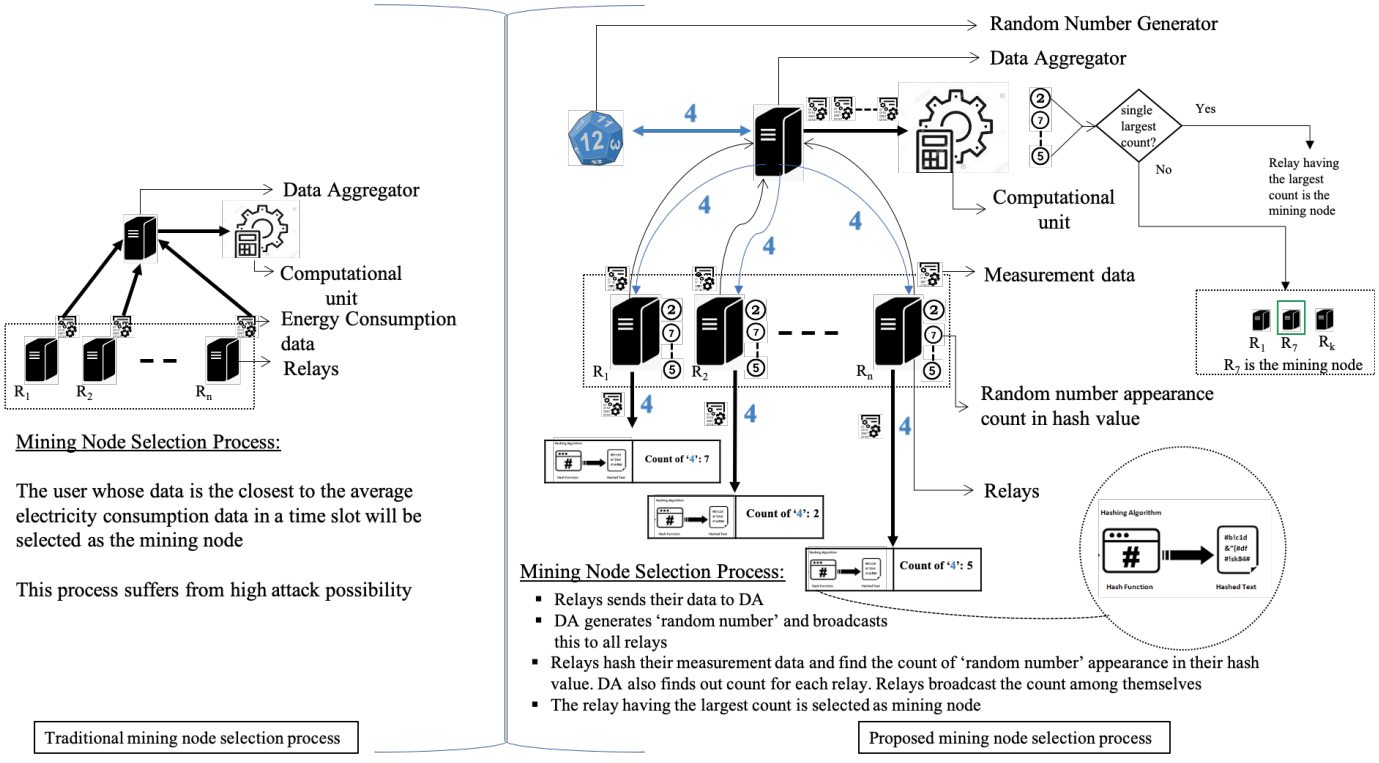


Fig. 3. Traditional vs proposed mining node selection procedure.

other. Meanwhile, the DA also hashes each measurement and finds out the same counts against each hash value. In this stage, every server including the DA has the same counts. At the final stage, all the servers cast their votes for the relay server having the largest count of random number appearance. If all the relay servers reach a consensus, the DA server selects the relay having the largest count as the mining node for that cycle. However, if multiple relay servers have the same largest counts or all of them have the count as zero (0), then the DA selects the mining node randomly following a cryptographically secure random selection algorithm. The functions $random.randint(start, end)$ and $random.choice(list)$ of the module $random$ from the python standard library has been used to generate random numbers and select random node in our simplified demonstration. However, for achieving more security and randomness, the proposed scheme can be extended to adopt CSPRNG (cryptographically secure pseudorandom number generator) algorithm. We have considered to include it in our future works.

Block creation, verification and addition: The working principle of the simplified demonstration of the scheme has been described in Fig. 4. Here, as soon as the mining node selection is completed, the DA sends the encrypted and aggregated measurement data to the mining node. The measurement data is then hashed by the mining node using the technique described under III-E1. During hashing operation, a Merkle tree structure has been followed. Then, the mining node creates a new block with the previous hash, index, timestamp, consumption value, current hash (root hash), and nonce. A typical block has been shown in Fig. 2. All the relay servers and the DA participate in the block verification process. During the verification process, the servers exchange responses several times. If all parties agree on the validity of the block, the DA sends the block addition request to all the servers. Finally, after the successful addition of the block to the blockchain, the DA sends the updated blockchain to the control center.

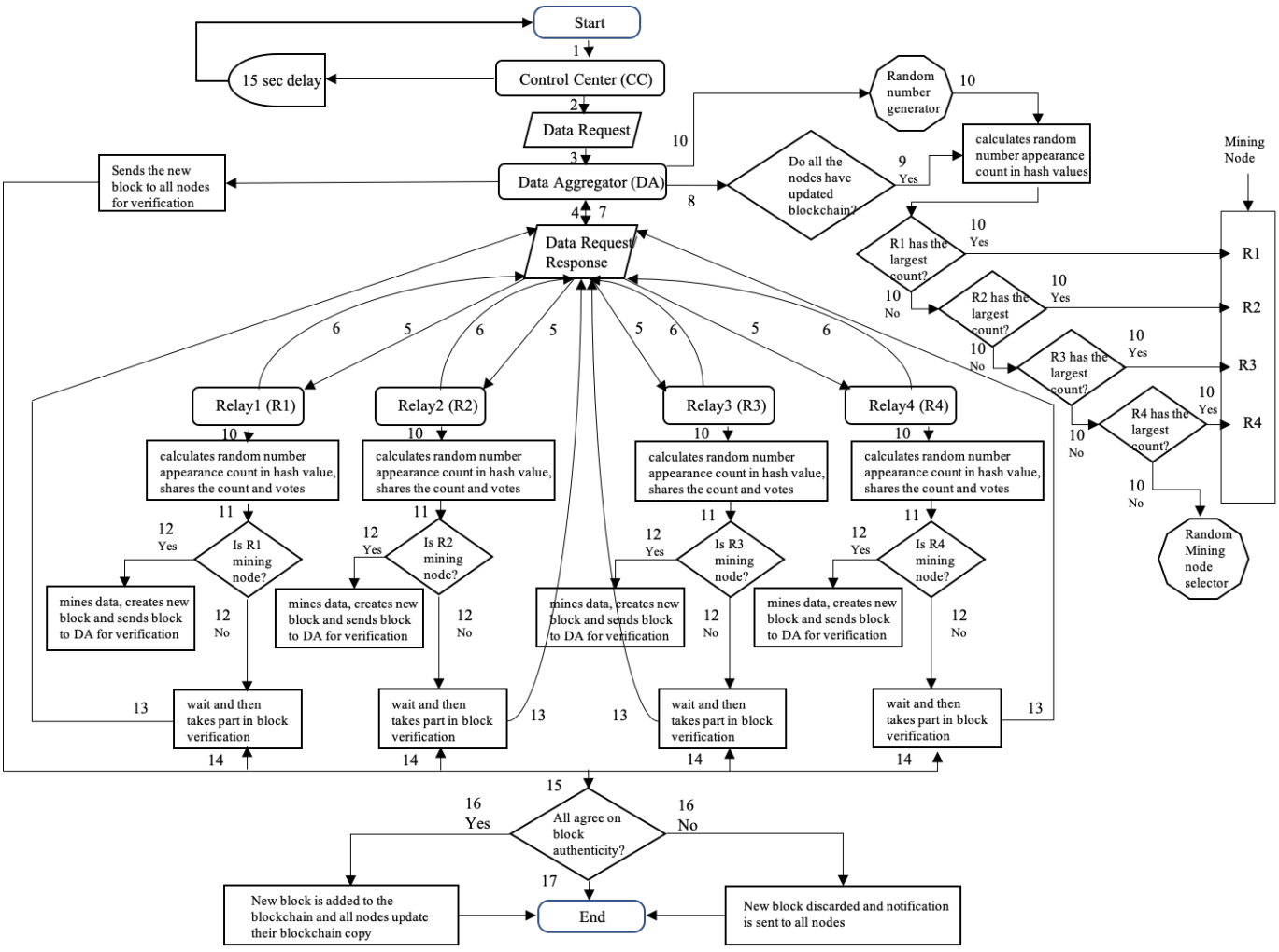


Fig. 4. Flowchart of the operational procedure of proposed scheme. The process follows the number sequence of the flowchart.

IV. Performance Evaluation

A. Experimental Setup

For experimental purposes, we set up a testbed in a virtual machine. Ubuntu 20.04 (64-bit) with a base memory of 8 GB and storage of 20 GB has been considered as the operating system. To build up servers/nodes, Mininet version 2.2.2 has been used. We have used Python 3.0 as the programming language. Moreover, We have used Wireshark (previously known as Ethereal) to capture the network packet (see Fig. 5). DNP3 runs over TCP and UDP. The Wireshark DNP dissector registers for TCP and UDP ports 20000 by default and checks whether the TCP segment it's handed has at least 2 bytes of data or not. The measurement data is coming from an IEEE 9-bus system [19].

B. Result and Discussion

As this is a no reward-no penalty based private and permissioned blockchain model for smart industries, miners/validators do not need to compete for mining opportunities. So, a complex mathematical problem-solving approach has been avoided which in turn saves energy that may incur due to huge computational overhead as it occurs in PoW. At the same time, the mining process needs to be random and fair which can also provide security to the data acquisition system. The proposed approach is a good fit in this context. It has the below advantages over them-

- One of the goals of the paper is to find out an efficient consensus mechanism for private blockchain which is suitable to implement in SCADA systems in the context of smart grids. So, incentive or penalty mechanisms have been avoided in the proposed scheme. This, in turn, solves the huge computational overhead or unfair selection issues of PoW or PoS mechanisms
- A level of fairness has been ensured as all the relay servers get the chance of voting. At the same time, the random number generator and the random node selector ensure the randomness of the mining node selection process
- As all the servers have the same counts, even if somehow, an adversary compromises a server, he/she cannot alter the count to become a mining node without revealing his/her identity. This provides security in mining node selection process

However, questions may arise regarding the complexity and computational time for this process. It is desired to be less time consuming, especially for the operations where near-real-time data is very important. Surely, selecting the mining node through a simple random process by the DA would reduce both the complexity and computational time. But, in that case, an adversary may easily become a mining node by only compromising the DA. So, a trade-off has been made between the computational time, complexity, and security.

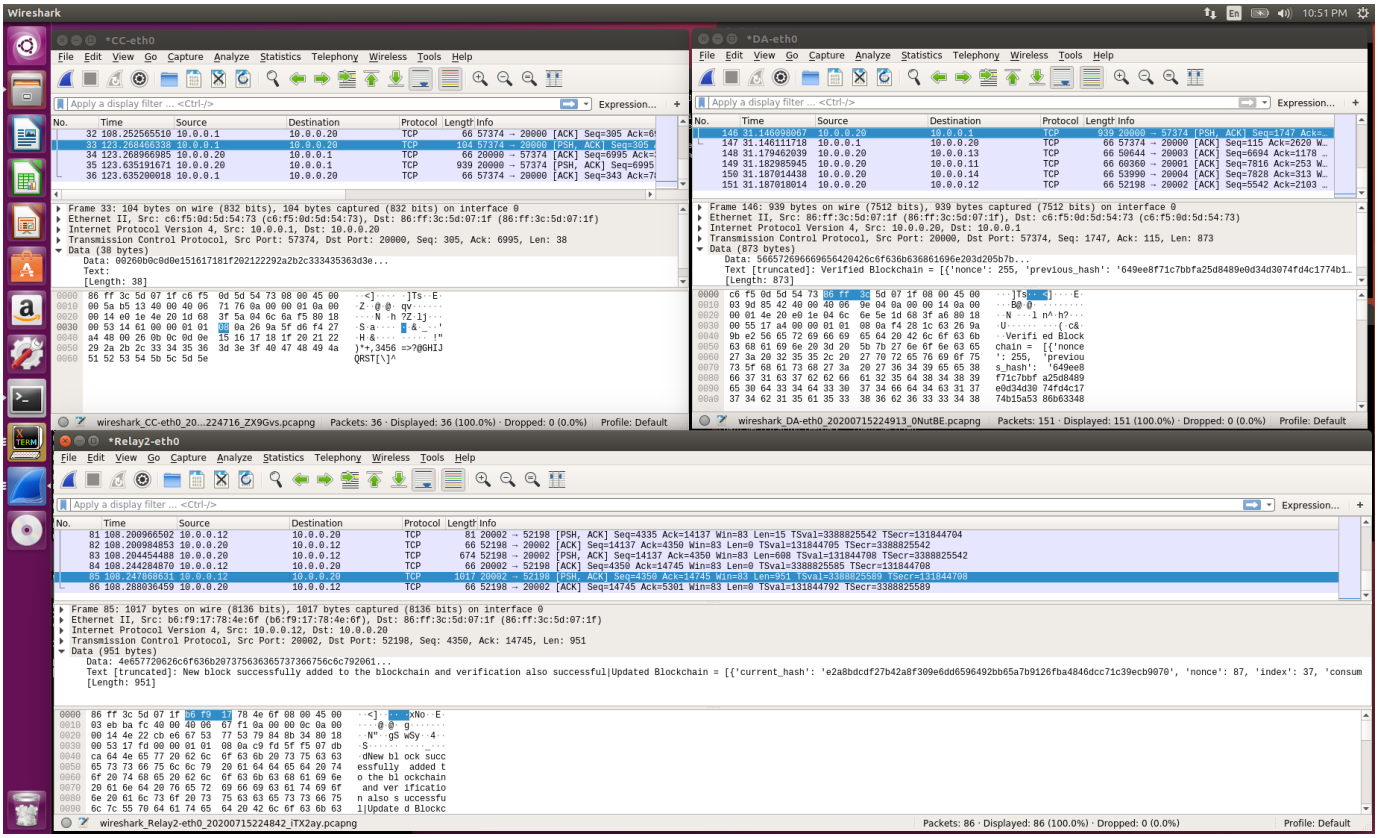


Fig. 5. Network packet analysis using Wireshark.

The simplified experiment shows that an entire operation cycle with four relay servers and four measurement data (one from each server) per block takes only 382 milliseconds. Mining node selection procedure takes around 75% of the processing time of a cycle (see Fig. 6).

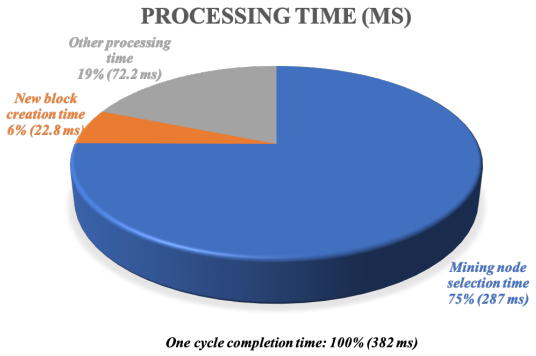


Fig. 6. Processing time of the prototype.

If the number of nodes increases, the processing time would also increase. One possible way to overcome this is to allow only a subset of total nodes, say for an example, K nodes to be eligible as a mining node among N nodes where $K < N$ and $K \subset N$. Then, at a certain point, the increment of mining node selection time would be halted even if the number of nodes increment goes on. Other possible measures can be taken to make the data acquisition process more secure and trustworthy, e.g., data encryption, data anonymization using pseudonyms, node identity anonymization. Moreover, faster authentication could be achieved by using a bloom filter. There is a very low possibility of hash collision during the hashing mechanism. Reference [6] proposed to decrease the probability of hash collision by increasing the array size of the bloom filter. There is a debate on using double-SHA-256 to overcome the hash collision but bitcoin is using it as $SHA256(SHA256(x))$ [20], where x is the block_header.

V. Conclusion

In this paper, a novel consensus mechanism named PoRCH with a personalized mining node selection process has been developed from a practical point of view for the SCADA systems in the field of smart grids and Industry 4.0. The scheme bridges the OT protocol, DNP3 to the blockchain technology while maintaining randomness and fairness of the process. The paper also presents a simplified and lightweight demonstration of the proposed scheme. The demonstration verifies the low computational overhead requirement of the scheme and reflects the characteristics of a blockchain model.

References

[1] M. Iansiti, Karim R. Lakhani, "The Truth About Blockchain", Harvard Business Review, pp. 118–127, Harvard University, 2017

[2] M. Mao, H. Xiao, "Blockchain-based Technology for Industrial Control System CyperSecurity", 2018 International Conference on Network, Communication, Computer Engineering (NCCE 2018), Chongqing, China, 2018, pp. 903-907, Atlantis Press, doi: 10.2991/ncce-18.2018.151.

[3] A. Bhattacharjee, S. Badsha, A. R. Shahid, H. Livani and S. Sengupta, "Block-Phasor: A Decentralized Blockchain Framework to Enhance Security of Synchrophasor," *2020 IEEE Kansas Power and Energy Conference (KPEC)*, Manhattan, KS, USA, 2020, pp. 1-6, doi: 10.1109/KPEC47870.2020.9167676.

[4] A. Islam and S. Y. Shin, "BUS: A Blockchain-Enabled Data Acquisition Scheme With the Assistance of UAV Swarm in Internet of Things," in *IEEE Access*, vol. 7, pp. 103231-103249, 2019, doi: 10.1109/ACCESS.2019.2930774.

[5] A. Maw, S. Adepu, and A. Mathur, "ICS-BlockOpS: Blockchain for operational data security in industrial control system," in *Pervasive and Mobile Computing*, vol. 59, pp. 101048, Oct. 2019, doi: 10.1016/j.pmcj.2019.101048.

[6] Z. Guan et al., "Privacy-Preserving and Efficient Aggregation Based on Blockchain for Power Grid Communications in Smart Communities," in *IEEE Communications Magazine*, vol. 56, no. 7, pp. 82-88, July 2018, doi: 10.1109/MCOM.2018.1700401.

[7] Z. Zheng, J. Pan and L. Cai, "Lightweight Blockchain Consensus Protocols for Vehicular Social Networks," in *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 5736-5748, June 2020, doi: 10.1109/TVT.2020.2974005.

[8] D. Puthal, S. Mohanty, V. P. Yanambaka, and E. Kougianos, "PoAh: A Novel Consensus Algorithm for Fast Scalable Private Blockchain for Large-scale IoT Frameworks,", *Preprints*, 2020.

[9] *"DNP3 Overview Page,"*. Accessed on: Jul 16, 2020. [Online]. Available: https://www.dnp.org/About/Overview-of-DNP3-Protocol

[10] I. Vakilinia, S. Vakilinia, S. Badsha, E. Arslan and S. Sengupta, "Pooling Approach for Task Allocation in the Blockchain Based Decentralized Storage Network," *2019 15th International Conference on Network and Service Management (CNSM)*, Halifax, NS, Canada, 2019, pp. 1-6, doi: 10.23919/CNSM46954.2019.9012719.

[11] S. Kudva, R. Norderhaug, S. Badsha, S. Sengupta and A. S. M. Kayes, "PEBERS: Practical Ethereum Blockchain based Efficient Ride Hailing Service," *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, Doha, Qatar, 2020, pp. 422-428, doi: 10.1109/ICIoT48696.2020.9089473.

[12] S. Badsha, I. Vakilinia and S. Sengupta, "BloCyNfo-Share: Blockchain based Cybersecurity Information Sharing with Fine Grained Access Control," *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, USA, 2020, pp. 0317-0323, doi: 10.1109/CCWC47524.2020.9031164.

[13] I. Vakilinia, S. Badsha and S. Sengupta, "Crowdfunding the Insurance of a Cyber-Product Using Blockchain," *2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, New York City, NY, USA, 2018, pp. 964-970, doi: 10.1109/UEMCON.2018.8796515.

[14] Y. Zhao, "Optimizing Hash Strategy to Avoid Birthday Attack," in *Journal of Physics Conference Series*, Chengdu, China, vol. 1486, pp. 32004, Dec. 2020, doi: 10.1088/1742-6596/1486/3/032004.

[15] C. S. Gao, "Classification with Hash Collision Networks." , *ProQuest Dissertations Publishing, UCLA*, 2018. Accessed on: Jul 16, 2020. [Online]. Available: https://escholarship.org/uc/item/7421j1pj.

[16] D. Ongaro, J. Ousterhout, "In Search of an Understandable Consensus Algorithm," in *Proc. of the 2014 USENIX Conference on USENIX Annual Technical Conference*, USENIX Association, Philadelphia, PA, USA, Jun. 2014.

[17] M. B. Mollah et al., "Blockchain for Future Smart Grid: A Comprehensive Survey," in *IEEE Internet of Things Journal*, doi: 10.1109/JIOT.2020.2993601.

[18] S. Kudva, S. Badsha, S. Sengupta, I. Khalil and A. Zomaya, "Towards secure and practical consensus for blockchain based VANET," in *Information Sciences*, vol. 545, pp. 170–187, Feb. 2021, doi: 10.1016/j.ins.2020.07.060.

[19] K. Loji, I. E. Davidson and R. Tiako, "Voltage Profile and Power Losses Analysis on a Modified IEEE 9-Bus System with PV Penetration at the Distribution Ends," *2019 Southern African Universities Power Engineering Conference/Robotics and Mechatronics/Pattern Recognition Association of South Africa (SAUPEC/RobMech/PRASA)*, Bloemfontein, South Africa, 2019, pp. 703-708, doi: 10.1109/RoboMech.2019.8704802.

[20] R. P. Naik, "Optimising the SHA256 Hashing Algorithm for Faster and More Efficient Bitcoin Mining 1," UCL, London, Sep. 2013.

Security of Cyber Physical Systems: Vulnerabilities, Attacks and Countermeasure

Faisal Alrefaei
Department of Electrical Engineering and Computer Science
Embry-Riddle Aeronautical University
Daytona Beach, FL , USA
alrefaef@my.erau.edu

Abdullah Alzahrani
Department of Electrical and Computer Engineering
Oakland University
Rochester, MI, USA
alzahrani2@oakland.edu

Houbing Song
Department of Electrical Engineering and Computer Science
Embry-Riddle Aeronautical University
Daytona Beach, FL, USA
houbing.song@erau.edu

Mohamed Zohdy
Department of Electrical and Computer Engineering
Oakland University
Rochester, MI, USA
zohdyma@oakland.edu

***Abstract*—Recently, cyber physical system (CPS) has become one of the most useful computer systems used in critical infrastructure. The CPS has been effectively applied in many areas such as power systems, transportation, and robotics. The main focus of this paper is to review integrating the physical components with cyberspaces, avoiding some security issues, and ensuring system stability. In this review, we address what the CPS is, its architecture and why security is essential. This paper introduces some issues that result because of the integration. Also, this paper will be discussing some solutions that are used to enhance the security and protection of CPS, assessing the vulnerabilities in the CPS, detecting attacks to CPS and countermeasures.**

***Keywords*—Cyber physical system, vulnerability assessment, cyber threats and attacks, security.**

I. Introduction

Improvement of computer and network technology has been bringing a lot of advantages to people in their daily lives and businesses [1]. It can be now remotely: perform banking transactions, toggle automobile functions, and home monitor visitors, lights, and temperature controls all through computer infrastructure communications. The improvement of processor capabilities and transmission of data and command signals over the internet has led to increasing amounts of data signals transmitted to fulfill intended requirements. Recently, an innovation in cyber physical computer systems (CPS) has emerged that integrates system components to facilitate information flow in a secure communication cyber channels between components and equipment [3]. The CPS has been deployed widely on a large scale protect system integrity and privacy. All physical components are controlled and monitored remotely. The CPS control system analyzes data and releases decisions to be executed by physical devices based on complex algorithms at the application layer.

A data breach occurring during storage, transmission, collection, and operations could be the reason for losing information integrity over the network [1] and [3]. Resilient systems must therefore be designed to avoid, prevent, and mitigate attacks to vulnerabilities with effective countermeasures against catastrophic consequences. The cyber physical system (CPS) can be defined as a system that integrates physical sensor and actuator components with computation and transmission procedures in cyberspace to control and monitor operations effectively and sufficiently [6]. The CPS has been deployed to service a wide range of applications including banking, robotics, medicine, energy, transportation, military, chemical processing, and other industry [4] and [5]. That integration has attracted the attention of governments and free enterprise to the benefits of CPS in safety, reliability, robustness, stability, high performance, and efficiency.

Topics that will be addressed by this paper are issues, weaknesses, and solutions in CPS layers architecture. This starts from physical components in the physical layer such as sensors, actuators, and radio frequency identification (RFID) tags to the application layer through the transmission layer.

II. Importance of Security and Reliability in CPS

The essential importance of cyber physical systems (CPSs) has been growing significantly in the many areas where CPSs have been successfully applied [3] such as in controlling chemical processes and in automating buildings. These successes have made CPS control system components and network vulnerable to real-time attacks [7]. For example, hackers could use control system connection channels to deny service causing time delay responses that in turn could cause severe system-wide problems to power grid infrastructure [3]. Where uninterrupted reliable transformation functioning over the network is required, signal encryption could be an effective solution to enhance security. However, encryption alone may not be sufficient to reduce risks of internal staff disclosure or timing delays caused by complex error-prone algorithms [3], widespread system overload with a rapidly growing number of interaction devices [8], and security challenges given the

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

increasing CPS vulnerability to common attacks and intrusions [9]. The CPS systems must be considered essential to enable innovation of better methodologies that will reach higher levels of security and protection.

The cyber physical systems (CPSs) have been designed to work with application subsystems that are distributed or could be relocated to geographies lacking the same levels of physical security as other locations [10 - 12]. All the components of CPS are connected by the network, so if or when the connection endpoint is compromised, it could become a backdoor to other components on the network that could lead to random failures [14] or failures due to the system misinterpreting system damage inputs and responding incorrectly [15]. Perhaps because of these possibilities automatic control system fault-tolerant techniques such as adding of intrusion protection and prevention or redundancy may be added to the system to recover from potential damages [12].

All systems including CPSs are not immune from some kinds of threats such as faults, errors, bugs, or system corruptions that can happen randomly and cannot be avoided because of unanticipated implementation of the software; or internal or external, and active and passive attacker threats [12]. External threats could also include attacks made to what are considered constrained resources such as sensors. System tasks being processed could be interrupted if anything wrong happens to these external sensors. Internal attacks often appear as inherent flaws that would likely come from human errors in the design or implementation of hardware or software. Most of the fault that could be existing are likely to be mistakes in execution or configuration such as the disruption of preprogrammed set points [12] and [13].

III. Security Threats in the CPS

The widespread availability of the internet today makes it is easy to connect all systems and components because of the versatility of the internet protocols [16]. In such systems, threats are often interpreted as unauthorized access to sensitive information or equipment [12] in order to influence the system to behave in a malicious manner [17]. Most attackers try to interfere with physical equipment in the physical layer. They target the monitoring and control equipment [12]. As a consequence, the recovery will likely be taking place to restore the function and control of the system [18].

Recently, the attacks on sensors in what is called the perception layer in the CPS became an interesting research topic. Researchers addressed this kind of attack and suggested: First, reverse engineering can capture the node of the network as a tool to initiate counterattacks within this layer. Second, the attackers can physically or internally attack the node. They can use brute force to destroy the nodes. Alternatively, they can modify the internal code by using secret keys to gain access to sensitive information on the node. Third, the attacker can drain the node energy by executing unnecessary communication [12].

Security threats are often related to communications over and within the network the network layer. When successful, the direction of the messages will change, and not arrive at the expected time. The attacker uses denial of service (DOS) through spoofing that reduces or eliminates network execution [12]. The five types of DOS attacks are: jamming to interfere with signals to a node or a group of nodes; collision alters the protocol enabling collisions; routing stops the routing of messages in multi hop communication [12]; flooding sends too many connections to disrupt nodes; wormhole disrupts the node by creating a wormhole, deadening a signal.

Few security threats that can happen in the control layer [12] so risks mainly occur in desynchronizing CPS timelines. Decisions are executed that cause significant damages. The information layer passes and measures operational commands. Attackers use this kind of data to eavesdrop and analyze information flow. One strategy is to select a date, then expose access to the data taps and data carried by the network. Another strategy is to use traffic analysis to monitor the system and look for data characteristics [12].

IV. Cyber Physical System Layers Architecture and It's Issues With Solution

The CPS is a combination of physical objects with cyberspace that is connected through a network in three layers as shown in **Fig.1**. Captured data is executed in the physical layer by performing commands received from the application layer above [6]. The network layer transfers all commands. Collected date is analyzed and processed in the application layer analyses and sent back as commands to the physical layer. In this section, the functions, issues, and weaknesses of the layers are expanded upon.

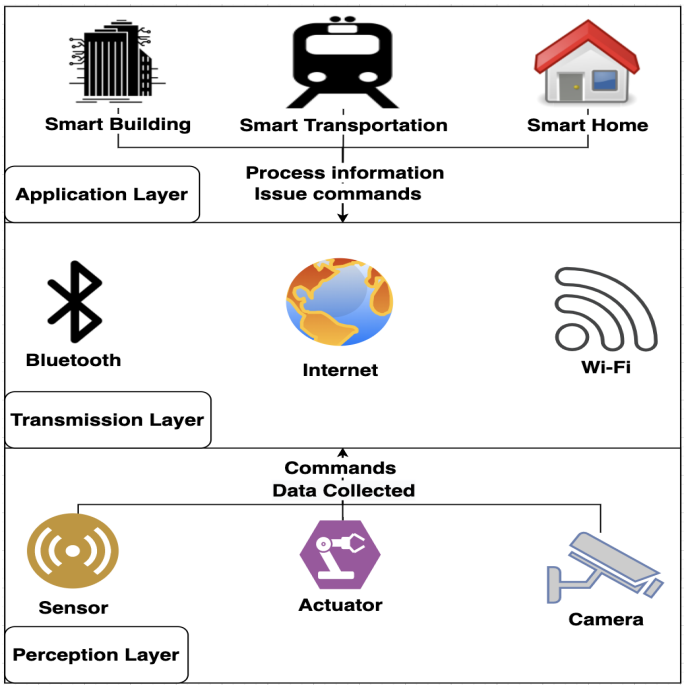


Fig. 1. The schematic diagram of three layers of cyber physical system

A. Application Layer

The application layer is the principal part of the CPS where smart environments are created. Functions such as the smart grid are housed together with other functions to be executed within intelligent environments. The first mission of this layer is to receive and analyze the execution control commands. All commands are sent from here using complex algorithms to perform their procedures upon the collected data [6]. Monitoring of the entire cyber physical system is another mission in this layer. The application layer monitors the behavior of the system and adjusts the behavior of the physical equipment to regulate what is preprogrammed for [6]. The storage system can also be placed within this layer to regulate any event that changes the behavior of the CPS. The most important function of this layer is securing the privacy of the data [7]. Any violations that happen to privacy here such as private data leaking, malicious code, and structured query language (SQL) injection attacks would be recognized as a threat to the entire system [19]. The policy of access control to this data including encryption and authentication mechanisms should be robust enough to ensure system security and privacy.

B. Network Layer

The primary function of network layer is transferring and interchanging large amounts of data and signal command control between the application and the physical layer [19]. The network layer is operated by the next generation of networks including the internet, local area network, Bluetooth, and Wi-Fi. There could be an issue at this layer when a large amount of data and transformation exceeds layer capacity. This can lead to congestion on the network and make it vulnerable to denial-of-service attacks (DOSs), distributed denial-of-service attacks (DDOSs), and black hole attacks. Deception attacks can lead to data loss integrity attack [20]. An effective way to resolve this issue is by considering encryption mechanisms [17].

C. Physical Layer

The physical layer is composed of embedded sensors, actuators, RFID tag readers, and mobile terminals [6]. These components connect the physical world to the physical layer. The central function of these components is either to monitor or track via gathering sound, heat, or light data and executing programmed commands. Once data is collected and transferred to an application, it is evaluated and analyzed enabling decision making. Tampering or modifying this data at this stage could trigger system wide catastrophic consequences. Consequently, protective secure communication code or cryptography is an imperative at this stage to provide enough privacy protection. One of the vulnerability concerns at this layer is that the controller keeps a secret key after decryption [21]. What is required is a form of encryption enabling the performance of calculations on encrypted data without decrypting the data first known as homomorphic encryption. Data has to be accurate and received encrypted in this way within the application [22].

V. Security Vulnerability and Assessments in CPS

Vulnerability assessment is a useful technique that identifies internal and external vulnerability presented in the system to enhance security. The majority of the proposed solutions have been previously focused on the prevention unauthorized users and operators from using or breaking into the system [21]. The vulnerability usually presents where there is an integration of systems with other control systems [22], [23] and [24]. Some totally vulnerable approaches fail to handle any probabilities of weakness within the system that consequently can be utilized by attackers to gain access to do damage to the system. As a result, researchers using the innovation vulnerability assessment approach have been performing ongoing research to create a method to support security in the system [25]. In aviation, for example, the security of CPS is critical, so all of the vulnerabilities must be identified and isolated immediately using this framework to assess the threats that present in all computer systems.

Another methodology has been created based on the idea that the vulnerability exists at the intersection of cyber, physical with cyber, and the human entities [26]. The main idea behind this methodology started with mapping the intersection of nodes in order of vulnerability of impact at each node for all manufacturing processes. The methodology was focused not only on malicious attacks, but also on any vulnerability that resulted from changes in manufacturing processes [26]. Unauthorized access to any control system is one of the main failures of security systems. This access can be detected by an intelligent mechanism or application that can detects and issue an alarm. Some attackers are unknown and are classified as a deceptive attack that is invisible to the system. Malicious attackers are always trying to reach the computer controllers to execute their malicious plans. Once the attacker has access to the system, they can make changes to the controller's set point by manipulating settings to drive the system to an undesirable outcome [22]. Hackers can inject false data to the controller or block the data that are sent to the controller. The attackers can also give false data to the operator in which some systems use a digital human machine interface (HMI). The HMI can be attacked, so that when false data is injected a wrong decision will be released. Also, attackers can disrupt the communication system by affecting the distributed control system and the supervisory control and data acquisition SADA where they depend entirely on the connection protocol [22].

VI. Attackers on the CPS

Attacks on the CPS lead to disclosure of data or possibly damage to the system using intelligent techniques by hackers. Hackers initiate their attacks based on timing factors because if time is not properly used, the impact could be ineffective. For example, false data injection (FDI), which is different from attacks made on traditional IT systems, is an attack that substitutes valid data with invalid data [23], [27] and [24]. The attack could be an active or passive attack [12]. Sophisticated CPS architectural design is divided into three main layers: application, network, and physical [6] and [23].

In the application layer, the attack targets access to and control over the control system and privacy of the user and operators. In the network layer, the attacks disclose or leak the data, altering or blocking the traffic to damage the system [15]. In the physical layer, there is an attack on the physical device or sensor that collects the data. Analyzing attacks are valuable to identify attackers' targets, formulate possible solutions and to build robust systems.

VII. How Intrusion Detection Works In CPS

Three techniques for intrusion detection include: misuse of signature-based, anomaly-based, and stateful protocol-based analysis. Misuse signature-based intrusion detection impacts how the signature defines any wrongdoing that happens so that the possible threats can be assessed. Anomaly-based intrusion detection is used for observed events to make sure the signature does not change. The stateful protocol analysis is used to check the activity executed for each protocol and stated events to examine changes [12]. Extensive memory is required to maintain signatures of some threat risks and it is difficult to identify the risks without knowing the exact name that was used for events. Signature detection intrusion is impossible to work well in CPS because the resource is constrained and the physical components in the CPS are widely separated. Stateful analysis is similar to anomaly-based detection. Activities of the protocol are referred to as normal behavior resources. The demand for the detection techniques in CPS is very high; thus, it is challenging to determine where the activities of the protocol will not be affected by threats. Anomaly detection is another name for the intrusion detection and is used when a detection pattern that interacts with a given data set isolates something that does not exist in normal behavior [12]. The benefit of using anomaly-based intrusion detection is that it does not use large amounts of memory. Flexibility and robustness are characteristics that distinguish this technique. It has been used as a suitable technique in CPS. The training data set collects the audit data regularly to produce a new training data set [12]. As a consequence, normal behavior for the system that is learned is produced when using the training data set. Therefore, when normal behavior deviates, it is assumed that the anomaly has detected everything required. This technique generates a report of the anomaly behavior, then sends it to the controller in order for any counter measures to take place to mitigate the attack.

VIII. Smart Home Security

The smart home is one of the most extensive applications concerning CPS [1] as shown in **Fig.2**. The smart home is operated by sensors and actuators that are connected through a network due to the architecture of the intelligent home. The centralized control is considered the main component due to its ability to control the home activities. All of these components are sensitive, and any damage or compromise that happens will cause major risk to the system. The owner may remotely do some management activities such as monitoring the temperature of the house by using the cameras to reduce the cost of utilities [1] and [2]. As a result of that, security has to be considered to avoid attacks that could lead to catastrophic consequences. A scheduling technique is utilized that

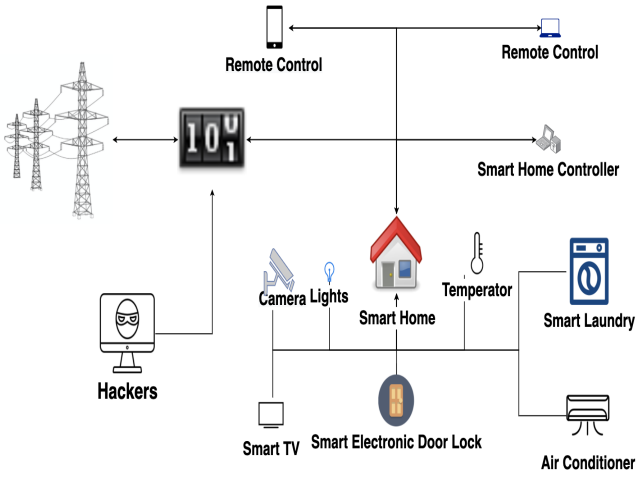


Fig. 2. The schematic diagram of the smart home security

assesses, monitors, and guides the electricity consumption in current and volts of household equipment to manage electricity consumption and loads. In this application there are two kinds of potential attacks on the smart home schedule [1]:

A. Pricing Cyberattack

Hackers manipulate guideline electricity pricing schedules to abnormally increase the power consumption data load of voltage and current to the smart house [1]. They reset pricing schedules to the lowest levels to mislead the pricing algorithm so that power loads are increased that in turn leads to increased electricity bills.

B. Energy Theft

The total electricity bill is calculated in the community based on each house's consumption of the last home [1] and [2]. The individual bill price and current is regulated based on the measurements used to calculate the community electricity bill. Hackers maliciously manipulate the flow of energy and the measurements on the smart meters to decrease loads of current. They manipulate the meter, which leads to a decrease in the electricity bill of one house, then leading directly to an increased bill at other houses.

IX. Secure Grid Power System

Secure grid [28] is a catchword that has been used in the last decade for having a power grid that is secure and reliable. This means having a wide-band secure power system with a circuit breaker (CB) at both the power stations and household premises, and remote terminal units (RTUs) with digital intelligence, combined with a precondition for having a grid that could be considered secure and reliable. Similar

information and IT security considerations will be encountered in the near future [29]. The protection of the power grid system can be an alternative direction as an extension of this work. Since the power grid system can be operated in two different modes such as grid-connected and stand-alone power, there are two different types of protections for each one that could be considered. The power grid protection system plays a key role in providing bidirectional protection between network components and device units. Thus, the inefficiency of the power grid protection approaches that are not restricted to an area with obvious edges can be addressed, and a flexible framework also can be found for protection scheme design of hybrid power grid applications. The power grid protection system can be modeled to have the capability to determine layers of fault discrimination to achieve minimization of electricity usage and reducing line losses. The future prospect of utilizing commercially available devices to protect such a system will be modeled in a future study. An issue may appear with very high impedance ground faults that may be tricky to fully detect. A knowledgeable researcher with a very good background in the electrical power system transmission and distribution will be required to carefully manage this situation. The fault protection and location are another part of this direction that will be based on the reliability of the circuit breaker. The protection of the power grid should be more robust, reliable and fast enough to protect the equipment on the system, reducing the error caused by the faults in the system can be modeled based on the number of protection zones and the speed of determining the fault on the line. Fast real-time communication is required for ensuring the protection of the power grid system [30]. For reliable and secure protection of the power grid operation, there should be a policy that ensures that it is undesirable to disconnect all power generators upon the occurrence of a fault. Advanced protection schemes are required for the purpose of detecting and locating the fault. The main protection schemes under grid connection and disconnection modes can be assumed on both modes of operation as shown in **Fig.3** and **4**.

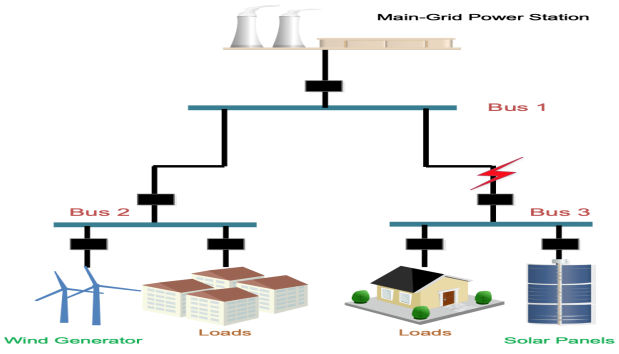


Fig. 3. A schematic diagram of the fault in grid-connected mode.

The philosophy of both modes of operation is nearly the same. However, there are some requirements that have to be considered based on how the relay switches are designed in some situations and what current value will be allowed on the circuit to pass through the switches for these two modes.

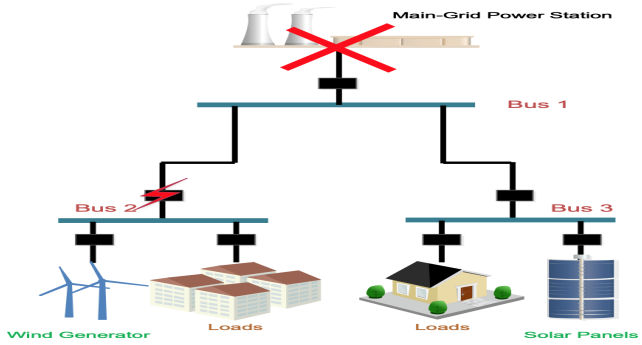


Fig. 4. A schematic diagram of the fault in Islanded-grid mode.

X. Conclusion

CPS is one of the next-generation technologies that will be used globally and will affect all aspects of our daily live. It is a technology that integrates physical entities with cyberspace. This integration leads to many users and devices interacting with massive amounts of generated data. These characteristics cause the technology to face many issues and challenges. Examples include vulnerability between entities, the security of data, command and security of physical components. That integration makes CPS different from traditional systems. Compared to CPS, conventional systems of security techniques are insufficient. Consequently, the CPS's security has been receiving ongoing research, innovating technologies and methods to enhance its reliability and safety.

References

[1] Wurm, J., Jin, Y., Liu, Y., Hu, S., Heffner, K., Rahman, F. and Tehranipoor, M., 2016. Introduction to cyber-physical system security: A cross-layer perspective. IEEE Transactions on Multi-Scale Computing Systems, 3(3), pp.215-227.

[2] Qiu, T., Zheng, K., Song, H., Han, M., & Kantarci, B. (2017). A local-optimization emergency scheduling scheme with self-recovery for a smart grid. IEEE Transactions on Industrial Informatics, 13(6), 3195-3205.

[3] Lyu, X., Ding, Y. and Yang, S.H., 2019. Safety and security risk assessment in cyber-physical systems. IET Cyber-Physical Systems: Theory & Applications, 4(3), pp.221-232.

[4] Sun, Y., & Song, H. (Eds.). (2017). Secure and trustworthy transportation cyber-physical systems. Springer Singapore.

[5] Ekedebe, N., Yu, W., Lu, C., Song, H., & Wan, Y. (2015). Securing transportation cyber-physical systems. In Securing Cyber-Physical Systems (pp. 163-196). CRC Press.

[6] Ashibani, Y. and Mahmoud, Q.H., 2017. Cyber physical systems security: Analysis, challenges and solutions. Computers & Security, 68, pp.81-97.

[7] Habibzadeh, H., Nussbaum, B.H., Anjomshoa, F., Kantarci, B. and Soyata, T., 2019. A survey on cybersecurity, data privacy, and policy issues in cyber-physical system deployments in smart cities. Sustainable Cities and Society, 50, p.101660.

[8] Alguliyev, R., Imamverdiyev, Y. and Sukhostat, L., 2018. Cyber-physical systems and their security issues. Computers in Industry, 100, pp.212-223.

[9] Lima, P.M., Carvalho, L.K. and Moreira, M.V., 2018. Detectable and undetectable network attack security of cyber-physical systems. IFAC-PapersOnLine, 51(7), pp.179-185.
[10] Song, H., Fink, G., & Jeschke, S. (2017). Security and Privacy in Cyber-Physical Systems. Wiley-IEEE Press.
[11] Song, H., Rawat, D. B., Jeschke, S., & Brecher, C. (Eds.). (2016). Cyber-physical systems: foundations, principles and applications. Morgan Kaufmann.
[12] Han, S., Xie, M., Chen, H.H. and Ling, Y., 2014. Intrusion detection in cyber-physical systems: Techniques and challenges. IEEE systems journal, 8(4), pp.1052-1062.
[13] Jiang, Y., Song, H., Yang, Y., Liu, H., Gu, M., Guan, Y., ... & Sha, L. (2018). Dependable model-driven development of CPS: From stateflow simulation to verified implementation. ACM Transactions on Cyber-Physical Systems, 3(1), 1-31.
[14] Walker-Roberts, S., Hammoudeh, M., Aldabbas, O., Aydin, M. and Dehghantanha, A., 2020. Threats on the horizon: Understanding security threats in the era of cyber-physical systems. The Journal of Supercomputing, 76(4), pp.2643-2664.
[15] An, Y. and Liu, D., 2019. Multivariate Gaussian-Based False Data Detection Against Cyber-Attacks. IEEE Access, 7, pp.119804-119812.
[16] Zegzhda, D.P., 2016. Sustainability as a criterion for information security in cyber-physical systems. Automatic control and computer sciences, 50(8), pp.813-819.
[17] Kim, N.Y., Rathore, S., Ryu, J.H., Park, J.H. and Park, J.H., 2018. A Survey on Cyber Physical System Security for IoT: Issues, Challenges, Threats, Solutions. Journal of Information Processing Systems, 14(6).
[18] Kong, F., Xu, M., Weimer, J., Sokolsky, O. and Lee, I., 2018, April. Cyber-physical system checkpointing and recovery. In 2018 ACM/IEEE 9th International Conference on Cyber-Physical Systems (ICCPS) (pp. 22-31). IEEE.
[19] Khalid, F., Rehman, S., & Shafique, M. (2020). Overview of security for smart cyber-physical systems. In Security of Cyber-Physical Systems (pp. 5-24). Springer, Cham.
[20] Gu, Q., Formby, D., Ji, S., Cam, H. and Beyah, R., 2018. Fingerprinting for cyber-physical system security: Device physics matters too. IEEE Security & Privacy, 16(5), pp.49-59.
[21] Kim, J., Lee, C., Shim, H., Cheon, J.H., Kim, A., Kim, M. and Song, Y., 2016. Encrypting controller using fully homomorphic encryption for security of cyber-physical systems. IFAC-PapersOnLine, 49(22), pp.175-180.
[22] Wardell, D.C., Mills, R.F., Peterson, G.L. and Oxley, M.N., 2016. A method for revealing and addressing security vulnerabilities in cyber-physical systems by modeling malicious agent interactions with formal verification. Procedia computer science.
[23] Butun, I., Österberg, P., & Song, H. (2019). Security of the Internet of Things: Vulnerabilities, attacks, and countermeasures. IEEE Communications Surveys & Tutorials, 22(1), 616-644.
[24] Dartmann, G., Song, H., & Schmeink, A. (Eds.). (2019). Big data analytics for cyber-physical systems: machine learning for the internet of things. Elsevier.
[25] Kumar, S.A. and Xu, B., 2017, June. Vulnerability assessment for security in aviation cyber-physical systems. In 2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud) (pp. 145-150). IEEE.
[26] DeSmit, Z., Elhabashy, A.E., Wells, L.J. and Camelio, J.A., 2016. Cyber-physical vulnerability assessment in manufacturing systems. Procedia manufacturing, 5, pp.1060-1074.
[27] Wolf, M. and Serpanos, D., 2017. Safety and security in cyber-physical systems and internet-of-things systems. Proceedings of the IEEE, 106(1), pp.9-20.
[28] Zheng, L., Gao, T., & Zhang, X. (2020). Security Protection and testing system for cyber-physical based smart power grid. In Proceedings of PURPLE MOUNTAIN FORUM 2019-International Forum on Smart Grid Protection and Control (pp. 847-857). Springer, Singapore.
[29] Song, H., Srinivasan, R., Sookoor, T., & Jeschke, S. (2017). Smart cities: foundations, principles, and applications. John Wiley & Sons.
[30] ALzahrani, A. M., & Zohdy, M. A. (2020). Real-Time Performance Assessment of Operating Photovoltaic (PV) Systems. Energy and Power Engineering, 12 (06), 339.

Modeling Static Cutting Forces And Stresses in the Face Milling Inserts

Richard Kyung
CRG-NJ
Cresskill, NJ, USA
richardkyung@choicerg.com

Jesang Yim
CRG-NJ
Cresskill, NJ, USA
jesangyim@choicerg.com

***Abstract*—This study has analyzed the mechanical processes that would lead to an optimal design of indexed milling tools that have WC inserts. This thesis deals with the finite element stress analysis at the rake face of milling inserts with three dimensional shape. It was idealized by simplifying cutting forces uniformly distributed loading and elastic behavior was assumed. Simulation was run on tungsten carbide specimens. Firstly, the shape of the insert cutting edges under the varying angles of the inserts were suggested. Then, an observed cutting force model was made. This permits analyzing the outcomes of the finite element stress on the cutting forces. A time-sensitive analysis was also made to observe the dynamic cutting forces and utility for the milling process. Afterwards, the models were put in a computer program to make the protocols of the models. The results show that there is a local maximum tensile stress on the rake face of the insert. It is also shown that there exist high shear stresses near the insert edge and that the normal stress decreases exponentially from a maximum value at the loading tool edge to minimum value at the other side of the insert. Results obtained are similar to those of orthogonal cutting but very much dependent on the shapes of the insert.**

***Key Words :* Face milling, Finite element stress, Three dimensional Insert, Cutting force, CNC, WC**

I. INTRODUCTION

In the past, the research on the milling insert was mainly focused on modeling the machinery and operations of common face milling cutters with different geometries, such as cylindrical, ball-end and tapered. Recently, several researchers have been utilizing analytical, numerical and experimental methods to analyze the mechanical efficiencies of various models of milling insert.

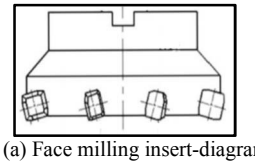
Organized tools appeared in the mid-20th century presenting innovative work in the study of inserted face milling cutters. Inserts represent financial advantages over other models of inserts since higher material removal rate (MRR) is potentially due to their high heat resistance. Analytical process models, that allow trustworthy predictions of cutting forces and machine-tool vibrations, provide the optimum process parameter selection, which is a problem before making cutting tests.

The face milling process is one of the most extensively used metal-cutting processes because of its high metal removal rate. The general forms of wear are the crater and flank wear in milling as in other cutting processes like turning. As cutting time increases the crater becomes deeper so that the resulting weakening of the insert edge can lead to breakage. Tool life is defined by specified values of these forms of wear of the flank and rake face.

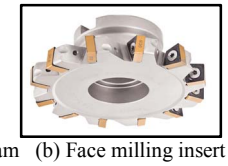
Main issues are to utilize the cutting forces and to avoid loud vibrations, as high and changing cutting force power generally indicate noises. This is the main decreasing factor for increasing MRR, as it may cause severe machining issues.

II. DATA AND RESULTS

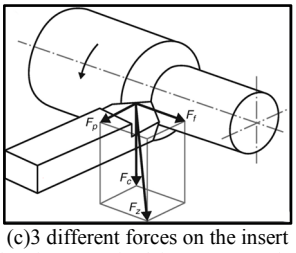
A. Face Milling Body and Forces on the Insert



(a) Face milling insert-diagram

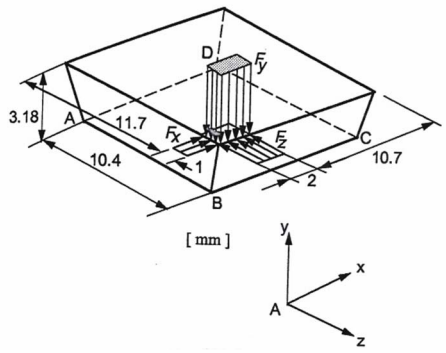


(b) Face milling insert



(c)3 different forces on the insert
(F_c:cutting force, F_f: feed force, F_p: passive force)

978-1-7281-9615-2/20/$31.00 ©2020 IEEE



(d) Sharp edged insert

Fig.1. Face milling insert(a)(b), 3 different forces on the insert(c) and assumed loading condition for the sharp edged insert(d)

For the finite element stress analysis for the insert shown in the Fig.1, the system was idealized by simplifying the cutting forces that are uniformly distributed along the surface and the elastic behavior was assumed.

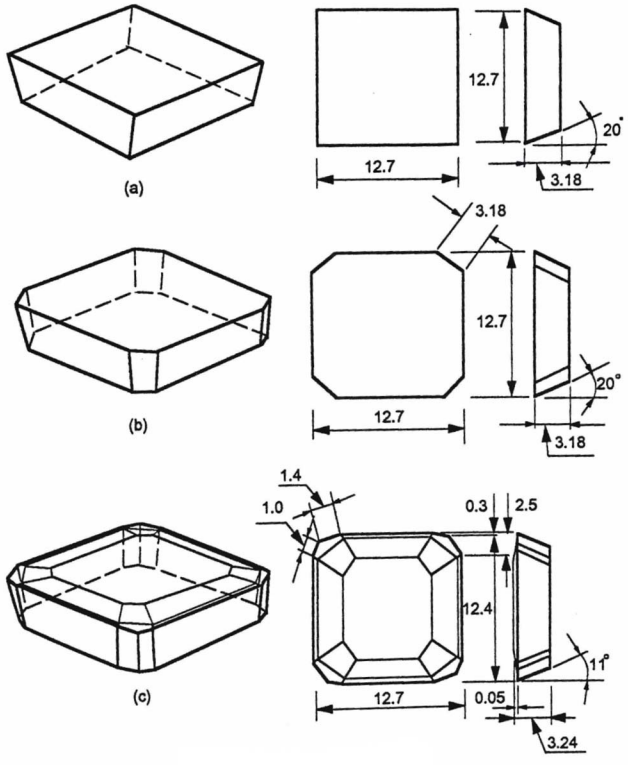


Fig.2. 3-type insert

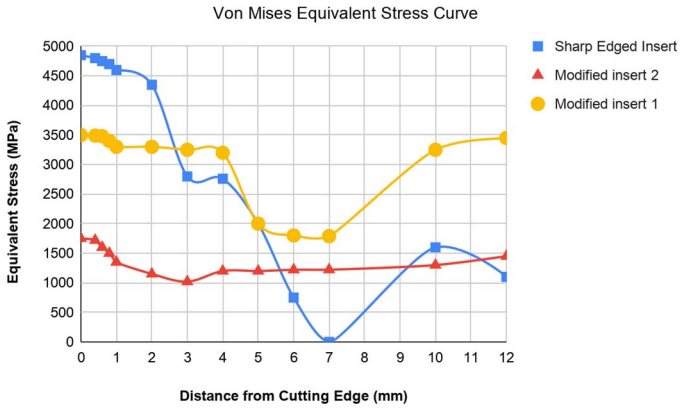


Fig.3. Equivalent stress distribution for the inserts

The results for the three different types of the inserts in the Fig. 2 show that there is a local maximum tensile stress on the rake face of the insert.

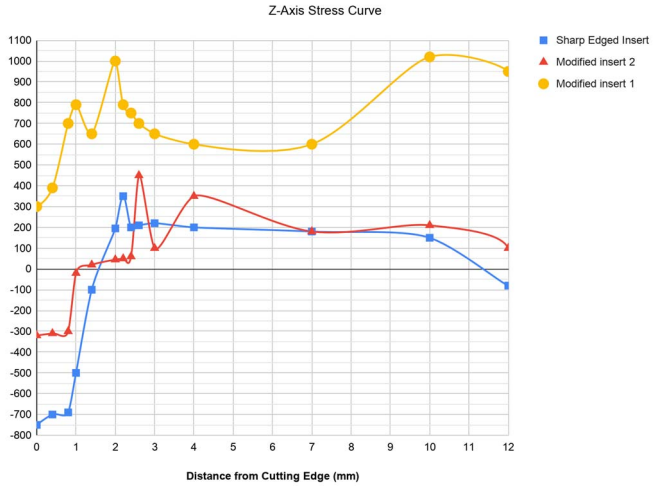


Fig.4. Stress distribution along z-axis for the inserts

It is also shown that there exist high shear stresses near the insert edge and that the normal stress decreases exponentially from a maximum value at the loading tool edge to minimum value at the other side of the insert.

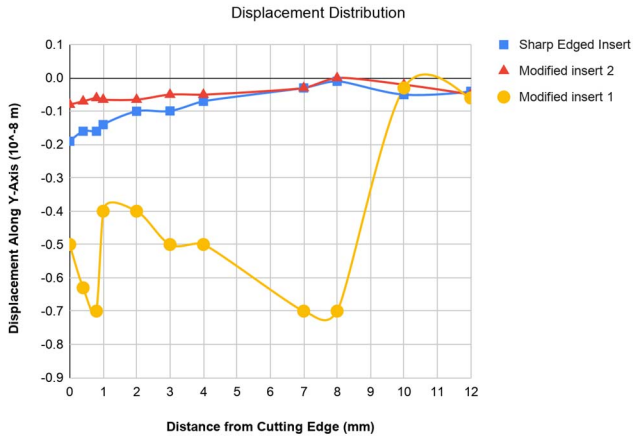
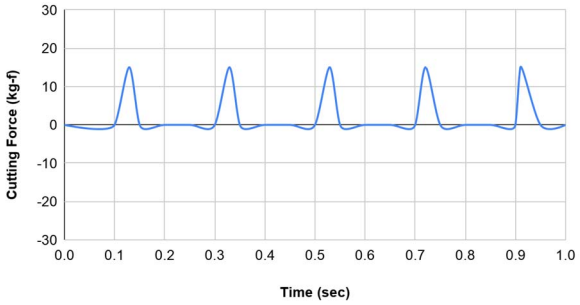


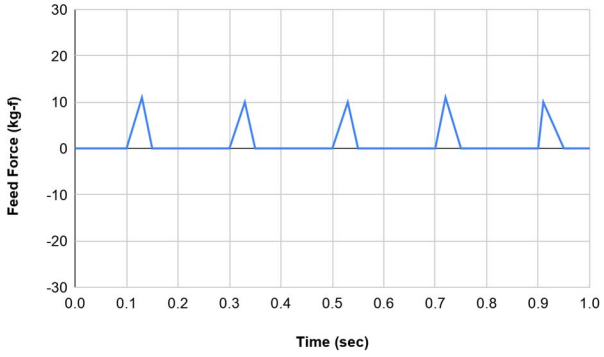
Fig.5. Displacement along the y-axis for the 3 inserts

There are three principal forces acting on the modified tool tip 2 during a turning process: The cutting or tangential force acts downward on the tool tip and the force is the largest among the three forces.

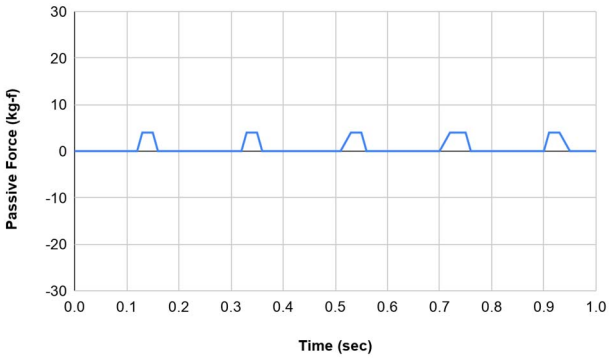


(a)

The feed force acts along the axis of the workpiece and perpendicular to the other two forces. The magnitude of the force is second largest among the three forces.



(b)



(c)

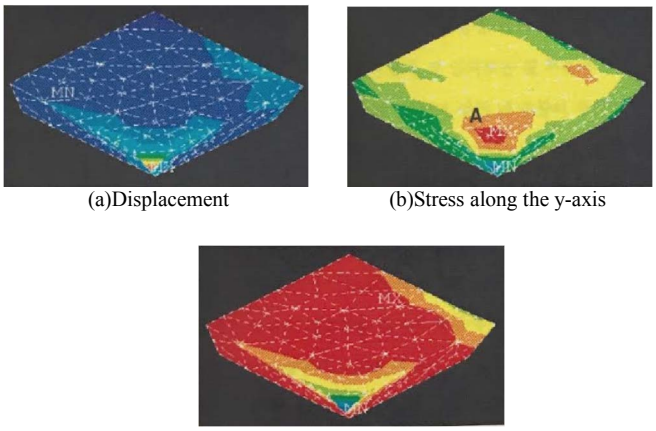
Fig.6. Cutting force data (modified insert 2)

The parallel force acts perpendicular to the axis of the workpiece and perpendicular to the other two forces. The magnitude of the force is the smallest among the three forces.

The results show that there is a local maximum tensile stress on the rake face of the insert. It is also shown that there exist high shear stresses near the insert edge and that the normal stress decreases exponentially from a maximum value at the loading tool edge to minimum value at the other side of the insert.

C. Stress and Displacement Distributions

To find stresses on the insert, forces along the x, y and z axes were assumed and FEM(Finite Element Method) was used to calculate the stresses. Deformations along the x, y and z axes in the inserts were also found.



(a)Displacement (b)Stress along the y-axis

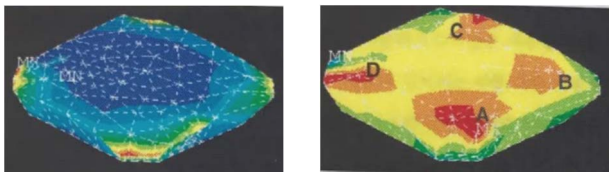
(b)Stress along the z-axis

Fig.7. Stress and displacement distribution for the sharp edged insert

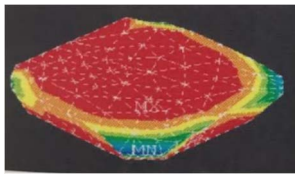
For the modified milling insert shown in the Fig. 8 shows:
- The least cutting forces
- Stresses(shear and normal) in the insert are smallest

The results for the three models(a, b and c) show local maximum tensile stress occurs on the rake face of the insert and high shear stresses occur near the insert edge.

Also normal stress decreases exponentially from a maximum value at the loading tool edge to minimum value at the other side of the insert.



(a)Displacement (b)Stress along the y-axis



(b)Stress along the z-axis

Fig.8. Stress and displacement distribution for the insert with modified edge (modified insert 2)

local maximum tensile stress was observed on the rake face of the insert.. It is also shown that there exist high shear stresses near the insert edge and that the normal stress decreases exponentially from a maximum value at the loading tool edge to minimum value at the other side of the insert.

IV. Discussion

Recently, researchers showed a uniform cutting force model for inserted cutters that can be utilized for various types of machining operations. Industry has developed new technology by using cutting tools with tungsten carbide inserts that are being operated in machining of super alloys. The tungsten carbide inserts cutting tools are more precise and therefore the use has been increased due to higher temperature resistance when compared to carbon inserts. However, the success of this process is dependent on the correct selection of process parameters.

Three different cutting forces were observed from cutting tests with a CNC milling machine. Then, the cutting force design was certified with cutting experiments measuring cutting forces and observing chips. After receiving modal testing, cutting forces and vibrations were assessed by means of the time sensitive model. A series of simulations were accomplished to conclude stability limits at certain operating conditions using the time domain model, and established a model.

The results show local maximum tensile stress occurs on the rake face of the insert and high shear stresses occur near the insert edge. Also normal stress decreases exponentially from a maximum value at the loading tool edge to minimum value at the other side of the insert. Results obtained are similar to those of orthogonal cutting but very much dependent on the shapes of the insert.

References

[1] Explain Cutting Tools Archived 2019-05-12 at the Wayback Machine, https://mechanicalsite.com, retrieved 2019-05-12.

[2] Stephenson, David A.; Agapiou, John S. (1997), Metal cutting theory and practice, Marcel Dekker, p. 164, ISBN 978-0-8247-9579-5.

[3] Usher, John T. (1896). The Modern Machinist (2nd ed.). N. W. Henley. Retrieved 2013-02-01.

[4] Practical treatise on milling and milling machines. Brown & Sharpe Manufacturing Company. 1914. Retrieved 2013-01-28.

[5] A treatise on milling and milling machines. Cincinnati, Ohio: Cincinnati Milling Machine Company. 1922. Retrieved 2013-01-28.

[6] Noble, David F. (1984), Forces of Production: A Social History of Industrial Automation, New York, New York, USA: Knopf, ISBN 978-0-394-51262-4, LCCN 83048867.

[7] Roe, Joseph Wickham (1916), English and American Tool Builders, New Haven, Connecticut: Yale University Press, LCCN 16011753. Reprinted by McGraw-Hill, New York and London, 1926 (LCCN 27-24075); and by Lindsay Publications, Inc., Bradley, Illinois, (ISBN 978-0-917914-73-7).

[8] Roe, Joseph Wickham (1916), English and American Tool Builders, New Haven, Connecticut: Yale University Press, LCCN 16011753. Reprinted by McGraw-Hill, New York and London, 1926 (LCCN 27-24075); and by Lindsay Publications, Inc., Bradley, Illinois, (ISBN 978-0-917914-73-7)..

Study on How Seniors and People with Disabilities are Adapting to New Technology

Richard Kyung
CRG-NJ(Choice Research Group)
Cresskill, NJ, USA
richardkyung@ChoiceRG.com

Ethan Park
CRG-NJ(Choice Research Group)
Cresskill, NJ, USA
ethanpark@ChoiceRG.com

***Abstract*– When closely examining data of technology users, there is an apparent and interesting shift in two specific groups—the elderly and disabled. This paper investigates the potential reasons for the data trends found in these two groups. Technology greatly impacts people with disabilities and how they navigate the world. Technological developments increasingly work to address the isolation that people with disabilities experience, and how they are often unable to engage with their communities. This research paper outlines the way technology has shaped social communication, information distribution, and day-to-day living among people with disabilities.**

The internet has transformed social communication and interaction. Previously, socially isolated people with disabilities have gained exposure to social environments through the internet and social media. The broad range of information available on the internet increases access to resources such as government services, health services, and social support. Assistive devices help people overcome their motor, sensory, or cognitive difficulties that may hinder them from performing daily functions. Although technology has proven to be effectively integrated in the lives of people with disabilities, many individuals with disabilities lack the access to technology such as a computer with internet connection. This paper examines how technology has benefited communication, information distribution, and day-to-day living in the disabled and aging community.

***Keywords* – Assistive devices, Technological developments, People with disabilities, Information distribution, Isolated people**

I. Introduction

The recent application of microfluidics in research involves examining cell growth and development in dynamic microenvironments [4]. The advantages of the microfluidic technology include miniaturization, parallelization, and integration. This research quantitatively studies the micro-fluid flow in organ-on-a-chip (OOC) microfluidic systems.

Defining necessary specifications builds the foundation in designing the structure of microfluidic channels. During this process, factors such as sample size, flow pattern, optimal flow rate, flow uniformity, and fabrication complexity must be considered.

Similar to a tubular network of blood vessels, a microfluidic channel that separates the flow into multiple channels is proposed. The goal of this particular system is to address the constraints on the microfluidic system.

Micro-fabrication provides a significant toolbox which helps process and manufacture the microfluidic channel. However, microfluidic systems are often costly due to the complexities that result from inconsistent channel depths and dimensions. This necessitates using sophisticated technology and establishing fabrication procedures. In order to fabricate the microfluidic channel, the proposed geometry must include and develop uniform cell deposition in proper devices.

The design of the proposed microfluidic channels has circular cross sections because of its low fabrication complexity and simple manufacturing procedure. With a Reynolds number less than 2000 the laminar flow requirement of such microfluidic channels is satisfied [1].

In this study, the channel diameter D is constant at nano-scaled dimensions of 4E-9m and 3E-9m. In addition, the cross-section of the channels will be circular: a few nanometers in diameter, and a few millimeters long. Multiple channels may be used in a device by either joining them (parallel) or having many disconnected ones (independent).

II. Technology in the Twenty-First Century

Though the term technology typically elicits conceptions of digitization, screens, and internet usage, at its core, technology simply defines the practice of combining skills and ideas towards a goal. In the twenty-first century, an era that solidified the place for computers and the internet in everyday life, a new type of technology emerged for quotidian household and recreational purposes. Technology has transcended its initial purpose for logistical streamlining and industrial use and has revolutionized the ways every individual communicates with others and perceives the world. This emergence of daily technology is more pronounced for individuals with disabilities and handicap; as the definition of technology is to create something to achieve a goal, every item meant to assist disabled individuals is an example of technology in and of itself. In the twenty-first century, technology for disabled individuals surpasses mere assistance, rather, it allows for an enhanced, multi-dimensional quality of life.

Technology acts and an aid and transmitter for the disabled who were previously oftentimes isolated from social environments. New advances in technology feeds information, allows for social interaction, aids and helps in disabling factors, acting as a liaison to the outside world through regular interaction. This puts disabled people in the relationship as opposed to outside of it. Now the diasbled can navigate more fluidly, engage in otherwise limited activities, and do things

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

that they once never thought was possible. For this, there are more options and more ways of becoming as independent as possible, offering new and growing opportunities to much needed populations.

Digital Disparity

Developments in technology have reshaped the life expectations of those with disabilities by enhancing their capacities and optimizing their quality of life. Technology has provided people with verbal disabilities a chance to communicate, with audial disabilities a chance to hear, with mobile restrictions a chance to explore. When people are more exposed to social environments, they have the ability to experience higher levels of social and personal growth, as well as the ability to make impactful contributions. Such technological advances are increasingly allowing people with physical disabilities to live a more participatory and stimulating lifestyle. Acting as much need aid for the disabled, technology has shown to improve the quality of their lives and decrease the symptoms of their disability.

As the previous data suggests, there is actually a large correlation between certain socioeconomic factors and use of technology. Unfortunately, many individuals with disabilities do not have computers with internet connection in the household, or lack access to resources that can provide such technology due to low income. As resources evolve and improve, the lack of accessibility creates a digital divide in societal systems. Understanding the relationships between social status and access to technology is essential in this research. Being able to afford technology offers easier access to its benefits.

People who live in neighborhoods of high average socioeconomic status have higher access to technology. People who live in neighborhoods of lower than average socioeconomic status may have limited access to technology. This creates a digital divide between subpopulations of people with disabilities. This research examines methods to eliminate digital disparity and provide technology for all individuals with disabilities, creating opportunity for advancement.

Cell Phone Usage

The emergence of cellphones is an example of a technological solution to pre existing problems facing the disabled and elderly population. Cell phones have revolutionized the way people live and its convenience has been shown to aid the handicapped. Cellphones can be used by the disabled to connect, share, and acquire needed help when necessary. For many of the disabled, having a cellphone is necessary to the improvement and fulfillment of their lives. With 91% of Americans owning a cell phone, cell phone ownership is high, even among the elderly population. Among seniors of different demographic backgrounds, 77% own cell phones, with significant annual increases in cell phone ownership. Figure 1 breaks down cell phone ownership percentages by demographic characteristics such as age group, education level, and household income.

TABLE I. CELL PHONE AND SMARTPHONE USAGE PERCENTAGES AMONG SENIORS

Ranges	Cell phone	Smartphone
Total for all 65+	77%	18%
Age		
65-69	84	29
70-74	84	21
75-79	72	10
80+	61	5
Education		
High school grad or less	70	10
Some college	80	19
College graduate	87	35
Household income		
<$30,000	67	8
$30,000-$49,999	83	15
$50,000-$74,999	88	28
$75,000+	92	42

Although cell phones are prevalent among the elderly population, smartphones are yet to be widely adopted. A mere 18% of seniors are smartphone users, with slower adoption rates as compared to cell phone usage. Furthermore, the percentage of smartphone ownership drops precipitously with increasing age. 29% of seniors aged 65-69 own smartphones, but only 5% of seniors aged 80 and above own smartphones.

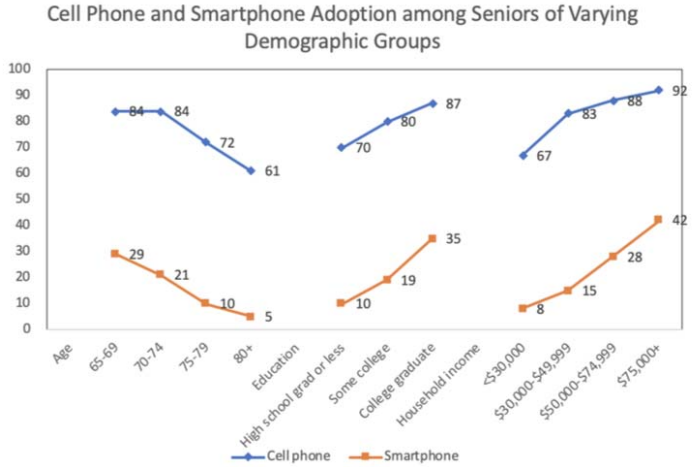


Figure 1. Table displays cell phone and smartphone usage percentages among seniors of various age groups, education levels, and household incomes.

On the contrary, the percentage of smartphone ownership increases in correlation with household income. Figure 1 shows smartphone usage by demographic characteristics such as age group, education level, and household income. The following graphics explore additional demographic factors such as education and income that affect technological ownership. As shown in Figure 3, e-book or tablet ownership

is higher among college graduates and those in higher income brackets. For example, 39% of seniors that earn $75,000+ own tablet computers, while only 8% of seniors that earn less than $30,000 own tablet computers.

TABLE I. E-BOOK READER AND TABLET COMPUTER USAGE PERCENTAGES AMONG SENIORS

	e-book reader	Tablet computer
Total for all 65+	18%	18%
Age		
65-69	23	23
70-74	19	18
75-79	18	20
80+	10	9
Education		
High school grad or less	12	11
Some college	19	19
College graduate	30	31
Household Income		
<$30,000	8	8
$30,000-$49,999	17	16
$50,000-$74,000	33	28
$75,000+	33	39

Compared to cell phone ownership, ownership of e-book readers and tablets is significantly lower in percentage. The percentage of people over 65 who own an e-book reader is 18%, and the percentage of people over 65 who own a tablet is also 18%.

Cell phones offer a wide variety of new and advanced options to people who are disabled, that is why they are such a useful tool in navigating disabilities. Many options and tools on cell phones are able to help disbaled people navigate the world better. For example, speak typing allows people with language barriers, reading problems, and vision issues to speak into the phone and receive answers to much needed information. When evr help is necessary cell phone sare equipped with emergency numbers and way of contact. Also, there are many other tools used on cell phones and tablets that are able to help disable people. This gives them a deepe opportunity to connect, communicate, and navigate the world around them.

The Internet and Global Connectivity

The invention of the internet has transformed modern society, changing the way people obtain information, interact with others, shop for products, and experience entertainment media. For people with disabilities, the internet provides an outlet of escape from isolation. Internet access has greatly improved the communication abilities of individuals with disabilities, in addition to their independence and self-sufficiency. Improved physical and mental health outcomes have also been reported as a result of internet use.

The internet encourages people with disabilities to be more visible and participate more fully in society. However, before people with disabilities can use the internet for communicating, learning, or working, they must be able to gain access to the internet. This often requires additional assistance by other people or by modified hardware and software. People with disabilities are less likely to live in households with computers that are connected to the internet and therefore less likely to be online. Even if they do have access to computers with internet connection, they may need the assistance of adaptive hardware and software. For example, a text reader is necessary for someone who is blind, and a speech to text program is required for someone with mobility issues that limit typing.

Developing standard, adaptive hardware and software is often difficult due to the broadly ranging levels of disability. An individual with hearing disabilities may solely require hearing aids, while others may require a sign language interpreter. Universally, closed captions for video and audio material are keys to successfully navigate and interact. Adaptive hardware and software are also necessary for individuals with vision disabilities. For people with low vision, a screen magnifier may suffice. However, for people with more significant visual loss, a screen reader that converts text to speech output may be required. In addition, Braille readers that convert text on the internet to Braille may be helpful.

People with physical disabilities often are unable to use a standard keyboard. In such cases, alternative keyboards with larger keys, different key arrangements, touchscreens, or speech recognition software using voice command may be warranted for internet navigation.

Some individuals may have poor hand or voice control such as in severe cases of stroke or Multiple Sclerosis patients. Assistive technologies exist and continue to emerge, allowing individuals to utilize their head, mouth, or eye movements as alternatives to a traditional mouse or keyboard.

For people with speech impairment, communication through discussion boards, blogs, and other social networking tools can be beneficial. Communication with speech output can be used when speech is necessary. For people with cognitive impairments such as learning disabilities, technology plays a large role in increasing or maintaining their independence. Tools such as a spell checker, grammar checker, work prediction, and voice recognition programs are valuable. Blogs and websites offer a wide array of information for the disabled. Discussion boards can clarify and use communication to interact.

Many people with disabilities use the internet for social interactions and cathartic support. We increasingly use the internet to develop and sustain relationships, as is exemplified by the dominance of Facebook, Twitter, and other social networks and dating sites. In addition, there are many online resources to assist individuals if they need information about their disabilities. Social networks offer a wide variety of options, opportunities, and advancements for the disbaled, including joining groups that they can interact with and learning much needed information about their particular disability. This allows the disabled to form relationships and learn more ways of aiding themselves.

Given such benefits of internet use, disabled community efforts to increase access to computers with internet connection should be a paramount objective. These social connection points can provide a source of support, and further support can be found on informative sites relating to disabilities to assist in understanding, awareness, and even solutions to minimize the impact of disabilities.

New developments in technology have also shown positive outcomes and been able to assist the elderly at in home care and facilities. This may include video-monitoring, remote health monitoring, fall detection, bed alerts, and other aids to assist in safety and quality of living (Miskelly, 2001). Cell phones and the internet have also been shown to be useful for communication, information, and sharing. Assistive devices have shown to promote independence and decrease symptoms of age disability (Spillman, 2005).

TABLE II. E-BOOK READER AND TABLET COMPUTER USAGE PERCENTAGES AMONG SENIORS

Year	% of all adults 18+ who go online	% of seniors who go online
2000	50	14
2002	59	19
2004	62	22
2006	74	35
2008	75	35
2010	79	41
2012	86	59

An increasing number of people are adopting technology, including the disabled. As shown in Figure 2, a high count of new and advanced services have been included to enhance people's lives. As indicated, innovative technology has been used by a growing population in order to aid in communication, information sharing, social interaction, all aspects that can aid the disabled in utilization.

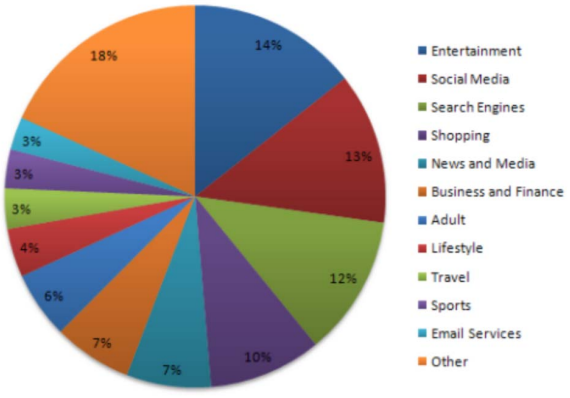


Figure 2. UK market share of internet visits

Nearly 13 percent of people use social media to form relationships and share information, whereas 14 percent use the internet for entertainment and 7 percent to get updates on the news. Technological aspects such as 12 percent search engines and 3 percent email services are useful tools for the disabled to learn, share, and communicate. This aspects have been shown to decrease symptoms of the disability and aid in the recovery and maintenance of a handicap.

Technology as a Way to Access Healthcare Information

In a study done in Pittsburgh, PA, participants attended an internet training course intended to introduce internet health services to seniors. At a mean age of 72, after a 5-week course participants were found to be more likely to use the internet as a source of medical information (Campbell, 2005). The study, which employed the Multidimensional Health Locus of Control Scale, the Lau, Hartman, and Ware Health Value Survey, as well as the Krantz Health Opinion Survey (Campbell, 2005).. However, the study failed to produce dramatic observed improvement in the shift from self-serviced medical care and tended to remain reliant on the physician-centered model. Nonetheless, seniors are responding to the internet use for more health-related information.

Another fascinating aspect to this study was the observed finding in demographic variability. Depending on demographic variables, participants varied in their behavioral decisions to seek health care online. This is in alignment with the other theories surrounding technology use and senior populations or disabled populations in that results are largely dependent on accessibility (Campbell, 2005). This study took special care to observe people from minority, low income, and limited educational backgrounds; with these demographics in mind, it is expected that access to top technology may be hindered. Nonetheless, technology proves to be an adaptable medium for seniors after some training and experience.

In addition, the researchers found that overall, giving seniors health autonomy via internet usage is a form of empowerment. The results found that an increased understanding of health allowed for seniors to have more self-worth and dignity surrounding matters of health and wellness which can ultimately lead to a higher quality of life and less disease (Campbell, 2005). Overall, this study has found that technology, even if minimal exposure to the internet, can benefit seniors greatly in their health journey.

III. CONCLUSION

Technological innovation can break the traditional barriers that exist for people with disabilities while limiting their exclusion and marginalization. Information and Communication Technology (ICT) devices and services can provide opportunities for people with disabilities to access lifelong education, skills development, and employment while facilitating communication and information distribution. According to a recent survey of 150 experts, websites and mobile devices can contribute to the social inclusion of people with disabilities. Text-to-speech, voice recognition, ability to change contrast and color schemes, touch and gesture input, and screen magnification features are already available

It is crucial to understand the disability divide as about 15 percent of the world's population lives with disabilities.

Discrimination exists in employing people with disabilities, which is illustrated by employment rates. The employment rate of people with disabilities are a third to half of the rate for those without disabilities. The unemployment rate among people with disabilities is as high as 80 to 90 percent in developed and developing economies. Unfortunately, lower education and lower productivity in a workplace results in a wage disparity among people with disabilities as compared to individuals without disabilities. Implementing community programs that help people with disabilities learn workplace technology will improve the high unemployment rate. This will promote new skills and increase learning abilities, making people with disabilities more accessible to the public.

However, there are several barriers and challenges that people with disabilities and seniors face when using new technologies. These include physical challenges to using technology due to disability, handicap, or chronic disease, skepticism about the benefits and adoption of technology, and difficulties learning to use new technologies. People with disabilities and seniors should increase their engagement in technology and the digital world. It is important to overcome the difficulties and learn how to use technology as there are many benefits to their mental health, daily life, and professional life. Communication channels also allow them to connect with more people and stay in touch with them. In addition, it allows them to find large amounts of health-related information beneficial to their health. It is imperative to address such issues and to find an effective way for more seniors and people with disabilities to use technology in their lives. Technology, when utilized, can be a benefit.

REFERENCES

1. Campbell, Robert J, and David A Nolfi. "Teaching Elderly Adults To Use The Internet To Access Health Care Information: Before-After Study". *Journal Of Medical Internet Research*, vol 7, no. 2, 2005, p. e19. *JMIR Publications Inc.*, doi:10.2196/jmir.7.2.e19. Accessed 20 July 2020.

2. Miskellly, F. (2001). Assistive technology in elderly. Department of Medicine for the Elderly. British Geriatrics Society.

3. Spillman, B. (2005). Assistive Devices Use Among The Elderly: Trends, Characteristics of Users, and implications for Modeling. The Urban Institute.

4. https://www.princeton.edu/futureofchildren/publications/docs/22_01_08.pdf
5. http://www.pewinternet.org/2014/04/03/attitudes-impacts-and-barriers-to-adoption/
6. http://pubdocs.worldbank.org/en/123481461249337484/WDR16-BP-Bridging-the-Disability-Divide-through-Digital-Technology-RAJA.pdf
7. https://www.ncbi.nlm.nih.gov/books/NBK97336/
8. https://gerontology.usc.edu/resources/articles/exploring-technologys-impact/
9. https://gerontology.usc.edu/resources/articles/how-technology-will-impact-aging-now-and-the-near-future/

Discovery of Influence between Processes Represented by Hidden Markov Models

Ritesh Ajoodha
School of Computer Science and Applied Mathematics
The University of the Witwatersrand, Johannesburg
South Africa
Email: ritesh.ajoodha@wits.ac.za

Benjamin Rosman
School of Computer Science and Applied Mathematics
The University of the Witwatersrand, Johannesburg
South Africa
Email: benjamin.rosman1@wits.ac.za

***Abstract*—Learning the underlying structure between processes is a common problem found in the sciences, however not much work is dedicated towards this problem. In this paper, we attempt to use the language of structure learning to address learning the dynamic influence network between partially observable processes represented as hidden Markov models (HMMs). The importance of learning an influence network is for knowledge discovery and to improve density estimation in temporal distributions. We learn the dynamic influence network, defined by this paper, by first learning the optimal distribution for each process using hidden Markov models, and thereafter apply redefined structure learning algorithms for temporal models to reveal influence relationships. This paper provides the following contributions: we (a) provide a definition of influence between stochastic processes represented by HMMs; and (b) expand on the conventional structure learning literature by providing a structure score and learning procedure to learn influence relationships between HMMs. We provide empirical evidence of the effectiveness of our method over several baselines.**

***Index Terms*—Learning influence networks, Structure learning, Hidden Markov models, Stochastic processes.**

I. Introduction

THE problem of describing the interaction or influence between stochastic processes has received little scrutiny in the current literature, despite its developing importance. Solving this complex problem has large implications for density estimation and knowledge discovery.

Usually, the individual structure of each stochastic process is ignored and all are merged into one big process which is modelled by some probabilistic temporal model. This approach undermines the explanatory importance of the relations between these processes [1]. The core of the issue is that we lose the underlying structure of the relationships between the processes which is essential to learn how one process influences the other.

In this paper, we provide a complete method for learning the dynamic influence network between processes represented by hidden Markov models (HMMs). This paper also explores the case when we are learning the influence relationship between partially observable processes. This is a significantly harder problem since the likelihood of the temporal model to the data has multiple optima which is induced from the missing samples [1]. Unfortunately, given that learning parameters from missing data is also a NP-hard problem, heuristic approaches are needed to solve for a suitable local optimum to the likelihood function of the parameters [1].

The application of this research is broad. Influence networks for stochastic processes can capture the complex relationships of how processes impact others. For example, we can learn the influence of traffic in a network of roads to determine how the traffic condition of a road will impact on another road. In educational data-mining we may want to determine the influence of participants in a lecture environment to encourage student success. We may wish to learn the influence between an IoT network [2]–[4]; influence between musical pieces [5], [6]; or influence between the skills of learners which impacts on their attrition [7]–[9].

The following contributions is made by this paper: (a) The concept of dynamic influence networks (DINs) which represents the influence (relationships) between partially observable stochastic processes. (b) The extension of the Bayesian information Criterion (BIC) for for dynamic models (d-BIC score). (c) The concept of a structural assemble which is able to relate dynamic models statistically. Finally (d), a greedy structure learning procedure for learning DINs between these HMMs.

II. Related Work

Numerous statistical procedures have been used to identify influence between variables [10]–[13]. These statistical procedures have been extended to the temporal environment to learn relationships between processes (variables over time). A significant contribution is the use of hidden Markov models (HMMs) which is defined as a set of parameters and conditional independence assumptions which together make up an acyclic structure between variables defined using factors [14]–[16]. The values in these factors are referred to as the parameters, and the list of conditional independence assumptions between variables are referred to as the structure of the model.

Learning the independence assertions of a dynamic Bayesian network can be used to make conditional independence inferences over time (density estimation) or to simply learn the relationships between variables (knowledge

978-1-7281-9615-2/20/$31.00

discovery). [17]–[20]. On the one hand leaning a sparse graph structure may have more generalisability for density estimation, but on the other hand, having a more dense graph can reveal unknown relationships for knowledge discovery. Care must be taken when considering for what purpose is the network required (more on this in the discussion) [16].

A successful approach to structure learning is using score-based structure learning [16], [21]. In score-based structure learning we develop a set of hypothesis structures which are evaluated using a score-based function that computes the likelihood of the data to the hypothesised structure. The likelihood is usually expressed as the information gain (mutual information) of the structure and parameters of the distribution to the data.

A search algorithm is then performed to identify the highest (possible) structure based on the structure score [22]–[25]. Viewing this problem as an optimisation problem allows us to adopt the already established literature on search methods in this super-exponential space to find the optimal structure given the data [26]–[29].

The structure of this section is as follows. In section II-A we introduce the well established BIC score which offers a way to trade-off the fit to data vs model complexity (the amount of independence assumptions between variables in the data). Finally, in section II-B we introduce a greedy search method to find the an optimal graph structure.

A. The BIC score

The BIC score models the structural fit to data verses the complexity of the conditional independence assumptions between variables. That is, the amount of independence assumptions made on the structure [30]. This makes it a popular choice for structure learning methods since the model complexity has a direct impact on the performance to do inference tasks. This is because the amount of conditional independence assumptions on a particular variable increases the factor size of that variable exponentially. The mathematical expression of the BIC score comprises of two terms: the first term models the fit to data; and the second term penalises the fit to data based on the complexity of the structure considered. The complete BIC score is as follows:

$$score_{BIC} = \ell(\hat{\theta}_{\mathcal{G}} : \mathcal{D}) - \frac{\log M}{2} DIM[\mathcal{G}],$$

where the count of instances is denoted by M and the count of independent parameters is denoted by $DIM[\mathcal{G}]$ in the Bayesian network. $\ell(\hat{\theta}_{\mathcal{G}} : \mathcal{D})$ represents the likelihood fit to the data.

The intuition of the Bayesian score is that as the amount of samples increase (ie. M) the score is willing to consider more complicated structures if enough evidence (samples, ie. M) is considered [31], [32]. The BIC core is particularly effective since the likelihood score (one without a penalty to complexity) will always prefer the most complicated network. However, the most complex networks also impose the risk of fragmentation, which is the exponential increase to the size of the factors caused by the increase of the in-degree of a variable. Penalty-based structure scores allows us to explore the opportunity to adopt more complicated structures if there is enough justification that the likelihood of the structure and parameters to the data is high-enough to compromise on the models speed to perform inference tasks caused by fragmentation.

There has been much contributions in the literature on the properties of the BIC score [30], [33], [34]. Key constitutions include a proof the it is consistent and is score equivalent which are necessary for efficient search procedures [35]–[37].

B. Learning General Graph-structured Networks

Since the search space for the optimal Bayesian structure is super-exponential, the difficulty of learning a graph structure for a Bayesian network is NP-hard. More specifically, for any $d \geq 2$, the problem of finding a structure with a maximum score with d parents is NP-hard [26]–[29]. See [38]–[40] for a detailed proof.

Despite this, there have been many contributions to learning an optimal structure. A key contribution is using heuristic search procedures to find an optimal acyclic graph structure [41]. These heuristic search procedures make use of search operators (changes to the graph structure) and a search algorithm (e.g. greedy search, best first search, simulated annealing e.t.c.) [42]. The intuition of this approach is to find an optimal acyclic structure by gradually improving the choice of the structure using the various search operators [43]–[47].

III. Influence between Hidden Markov Models

In this paper we consider a structure learning procedure which evaluates candidate dynamic influence networks (DINs) using scoring metrics. We provide evidence for the effectiveness of our structure learning procedure over the standard benchmarks selected.

An overview of the proposed structure learning procedure is given by the below instructions relating to Figure 1.

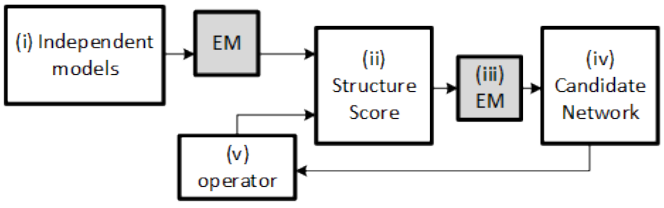


Fig. 1. An overview of the proposed structure learning procedure in this paper.

(i) The stochastic processes are given as input. The parameters for a HMM is learned for each stochastic process. This is the input.

(ii) A structure is imposed between the HMMs (using a relation function called an assemble). This gives us a dynamic influence network (DIN). The observable parameters are relearned in the model. The structure score for the DIN is computed.

(iii) Expectation maximisation is performed to learn the latent parameters of the DIN.
(iv) A candidate DIN is presented as output.
(v) The resulting DIN is evaluated and the score is recorded. If the score converges or a threshold is reached then the learning procedure is terminated. If not, we apply a structural operator (edge addition, deletion, reversal) and move back to step (iii).

Figure 2 illustrates an examples of a DIN between a set of HMMs. Each HMM is represented as a node in the acyclic graph structure and is denoted as a tuple, $\langle \mathcal{H}_0^i, \mathcal{H}_{\rightarrow}^i \rangle$, where $\mathcal{H}_0^i$ is the starting state of the HMM, and $\mathcal{H}_{\rightarrow}^i$ is the unrolled state for HMM i. A DIN structure is the output of the structure learning procedure.

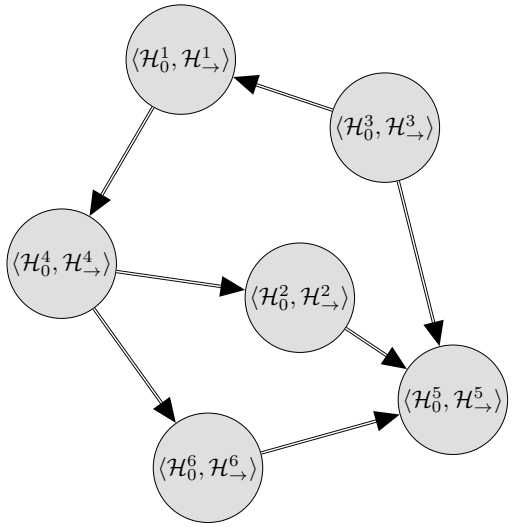


Fig. 2. A dynamic influence network (DIN) whose nodes represent six HMMs. Each HMM is represented as $\langle \mathcal{H}_0^i, \mathcal{H}_{\rightarrow}^i \rangle$, where $\mathcal{H}_0^i$ is the initial network and $\mathcal{H}_{\rightarrow}^i$ is the unrolled network. The double edges between each network represents the structure assemble (subsection III-B).

We will begin by providing a brief introduction to the hidden Markov model (HMM) which is used to represent the stochastic processes. An HMM is a dynamic Bayesian network (DBN). The likelihood function (mutual information) for a HMM, as illustrated by Figure 3, decomposes as:

$$L(\Theta : X^{0:T}, O^{0:T}) = \prod_{i,j} \theta_{X^i \rightarrow X^j}^{M[X^i \rightarrow X^j]} \prod_{i,k} \theta_{O^k|X^i}^{M[X^i, O^k]},$$

where the parameters correspond to the observable value k in the state i to the exponent of the number of times we observe both X^i and O^k. We will often refer to an HMM as a tuple as we did in Figure 2. [16], [48]–[50] provide excellent introduction to DBNs and HMMs.

In the section III-A we will extend the current literature of structure scores for Bayesian networks to scores for DINs; and finally, in section III-B we will introduce the notion of an assemble to relate HMMs in our DIN.

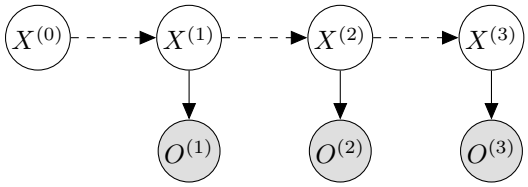


Fig. 3. An illustration of a hidden Markov model (HMM) with 4 time-slices. The dotted lines indicate the inter-time-slice edges for the persistent variable $X^{(t)}$. The solid line indicate the intra-time-slice edges for each respective time-slice.

A. Structure Scores for DINs

In Step (ii) of Figure 1, we needed to calculate the score of the influence structure. In this paper we adapt the celebrated BIC score to a dynamic BIC (d-BIC) for our dynamic influence networks. The d-BIC score make the same trade-off between model complexity and fit to the data, only the d-BIC can be applied to dynamic networks.

The d-BIC score is as follows:

$$\text{score}_{BIC}(\mathcal{H}_0 : \mathcal{D}) = M \sum_{k=1}^{K} (\sum_{t=1}^{T} (\sum_{i=1}^{N} (\mathbf{I}_{\hat{P}}(X_i^{\langle \mathcal{H}_0^k, \mathcal{H}_{\rightarrow}^k \rangle^{(t)}}; \mathbf{Pa}_{X_i^{\langle \mathcal{H}_0^k, \mathcal{H}_{\rightarrow}^k \rangle^{(t)}}}^{\mathcal{G}}))) - \frac{\log M}{c} DIM[\mathcal{G}],$$

where the amount of samples is given by M; the amount of dependency models is given by K; the amount of time-slices is given by T for any dependency model; the amount of variables in each time-slice is given by N; $\mathbf{I}_{\hat{P}}$ denotes the information gain in terms of the empirical distribution; and $DIM[\mathcal{G}]$ is the amount of independent parameters in the entire DIN.

The d-BIC score is designed to exchange the complexity of the DIN, $\frac{\log M}{c} DIM[\mathcal{G}]$, for the fit to the data, $\mathcal{D}$. As the amount of samples increases, the information gain term grows linearly, and the model complexity part grows logarithmically. The intuition of the d-BIC score is that we will be willing to consider more complicated structures, if we have more data that justifies the need for a more complex structure (i.e. more conditional independence assumptions).

B. Structure Assembles

Choosing the set of parent variables in a DIN establishes the notion of a structural assemble. A structural assemble is a template which relates temporal models. The structural assemble defines the parent sets for variables to construct a DIN. More specifically, the assemble relation is defined as follows:

Consider a family of hidden Markov models (H), where $\langle H_0^0, H_{\rightarrow}^0 \rangle$ represents the child with the parent set $\mathbf{Pa}_{\langle H_0^0, H_{\rightarrow}^0 \rangle}^{\mathcal{G}} = \{\langle H_0^1, H_{\rightarrow}^1 \rangle, \ldots, \langle H_0^k, H_{\rightarrow}^k \rangle\}$. Further assume that $\mathcal{I}(\langle H_0^j, H_{\rightarrow}^j \rangle)$ is the same for all $j = 0, \ldots, k$. Then the *delayed* dynamic influence network, denoted by $\langle \mathcal{A}_0, \mathcal{A}_{\rightarrow} \rangle$, will satisfy all the independence assumptions in

$\mathcal{I}(\langle H_0^i, H_{\rightarrow}^i \rangle)\ \forall\ i = 0, \ldots, k$. In addition, $\forall\ j$ and $\forall\ t$, $\langle \mathcal{A}_0, \mathcal{A}_{\rightarrow} \rangle^{(t)}$ also satisfies the following independence assumptions for each hidden or latent variable denoted L_i and some $t > \alpha \in \mathbb{Z}^+$:

$$\forall\ L_i^{\langle H_0^0, H_{\rightarrow}^0 \rangle^{(t)}} : (L_i^{\langle H_0^0, H_{\rightarrow}^0 \rangle^{(t)}} \perp\!\!\!\perp \text{NonDescendants}_{L_i^{\langle H_0^0, H_{\rightarrow}^0 \rangle^{(t)}}} | L_i^{\langle H_0^k, H_{\rightarrow}^k \rangle^{(t)}}, L_i^{\langle H_0^k, H_{\rightarrow}^k \rangle^{(t)-1}}, \ldots, L_i^{\langle H_0^k, H_{\rightarrow}^k \rangle^{(t)-\alpha}}, Pa_{L_i}^{\langle H_0^0, H_{\rightarrow}^0 \rangle^{(t)}}).$$

The assemble is an expressive representation to capture influence relationships that persist through time between temporal models in this case HMMs. However, the choice of α is important since choosing a large α will render many dependencies on variables. This causes a fragmentation bottleneck which causes a larger computational burden for learning and inference tasks.

To illustrate an example of using an assemble relation between two HMMs, $\langle \mathcal{A}_0, \mathcal{A}_{\rightarrow} \rangle$ and $\langle \mathcal{B}_0, \mathcal{B}_{\rightarrow} \rangle$, consider Figure 4. Figure 4 unrolls two HMMs, $\langle \mathcal{A}_0, \mathcal{A}_{\rightarrow} \rangle$ and $\langle \mathcal{B}_0, \mathcal{B}_{\rightarrow} \rangle$, using a structural assemble with $\alpha = 0$.

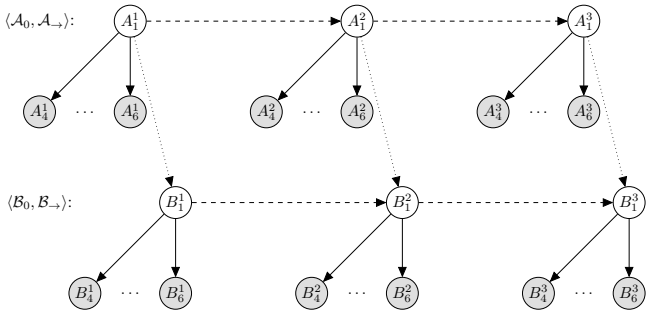


Fig. 4. Two unrolled HMMs, $\langle \mathcal{A}_0, \mathcal{A}_{\rightarrow} \rangle$ and $\langle \mathcal{B}_0, \mathcal{B}_{\rightarrow} \rangle$, as represented with 3 time-slices. The HMMs are connected by a structural assemble with $\alpha = 0$.

C. Structure Search

At this point we have the following well-defined optimisation problem:

1) A training set $\mathcal{H}_{\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}} = \{\mathcal{H}_{\langle H_0^1, H_{\rightarrow}^1 \rangle}, \ldots, \mathcal{H}_{\langle H_0^k, H_{\rightarrow}^k \rangle}\}$, where $\mathcal{H}_{\langle H_0^i, H_{\rightarrow}^i \rangle} = \{\xi_1, \ldots, \xi_M\}$ is a set of M instances from underlying ground-truth HMM $\langle H_0^i, H_{\rightarrow}^i \rangle$;
2) a structure score: score$(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle : \mathcal{H}_{\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}})$;
3) and, finally, we have an array of L distinct candidate structures, $\mathcal{G} = \{\mathcal{G}^1, \ldots, \mathcal{G}^L\}$, where each structure $\mathcal{G}^l$ represents a unique list of condition independence assertions $\mathcal{I}(\mathcal{G}) = \mathcal{I}(\mathcal{G}^{\mathbb{I}} \cup \mathcal{G}^{\mathbf{B}})$.

Our objective of this optimisation problem is to output the DIN which produces the maximum score. We present the following influence structure learning procedure in Algorithm 1, where $\mathcal{S} = \{\mathcal{S}_{\rightarrow}^1, \ldots, \mathcal{S}_{\rightarrow}^P\}$ represents the set of stochastic processes; $assemble$, is the option of the parameters for a structure assemble; and $score$, which is the selected scoring function used by the search procedure.

Algorithm 1 Influence structure search

1: **procedure** StrucSearch($\mathcal{S} = \{\mathcal{S}_{\rightarrow}^1, \ldots, \mathcal{S}_{\rightarrow}^P\}$, $assemble$, $score$)
2: for each process we learn a temporal model ($H = \{\langle H_0^1, H_{\rightarrow}^1 \rangle, \ldots, \langle H_0^P, H_{\rightarrow}^P \rangle\}$)
3: Using the models in H we generate a search space (ie. $\mathbf{G} = \{\mathcal{G}_1, \ldots, \mathcal{G}_n\}$)
4: Find the structure $\mathcal{G}_i$ which produces the highest $score$ (w.r.t. $assemble$) in $\mathbf{G}$
5: **return** $\mathcal{G}_i$
6: **end procedure**

The dynamic influence network, $\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}$, holds a distribution between a set of HMMs, denoted $\langle H_0^1, H_{\rightarrow}^1 \rangle$, $\ldots$, $\langle H_0^k, H_{\rightarrow}^k \rangle$, with the conditional independence assumptions listed by $\mathcal{I}(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}})$. We further assume that $P^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}})$ is induced by another model, $\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}})$, we will refer to this model as the underlying ground-truth model. The model is evaluated by recovering the set of local independence assertions in $\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}})$, denoted $\mathcal{I}(\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}))$, by only observing $\mathcal{H}_{\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}}$. This structure learning procedure is referred to by this paper as the greedy structure search (GESS).

D. Computational Complexity and Savings

The overall computational complexity of the above structure search algorithm is given by [1]. In order to allow for notable computational savings we suggest using a cache to store sufficient statistics and using a max priority queue (implemented using heaps) to arrange contending structures using their scores as keys. Random restarts and Tabu lists are also used to manage the structure search procedure.

IV. Empirical Results

This sections presents the performance of modelling influence between partially observable stochastic processes represented by HMMs using DINs. We evaluate the performance of our model aside several benchmarks.

The experimental setup is as follows. We constructed a ground-truth DIN which was used to sample sequential data. To simulate a partially observed process, several variables were removed from the sequential data sample. Algorithm 1 was used to learn candidate networks. Several variations of the algorithm was also used, such as using the d-AIC (dynamic Akaike Information Criterion) score instead of the d-BIC; using prior knowledge of the ground-truth structure such as the maximum in-degree used in the generative distribution; using a random structure; and even using no structure.

The parameters for the ground-truth DIN distribution is summarised by Table I. The ground-truth DIN distribution described the influences between 10 processes, each represented using HMMs with 5 time-slices, 2 hidden layers, 5 observable variables and 3 latent variables per time-slice. Each variable could take 3 discrete values. The overall ground-truth DIN had a max in-degree of 2 for any variable; and finally, the number of conditional independence assumptions (CIA) between processes was limited to 15.

TABLE I
A TABLE SUMMARISING THE PARAMETERS FOR THE GROUND-TRUTH DIN DISTRIBUTION.

Ground-truth DIN Distribution	
No. HMMs	10
Random variable values	3
No. time-slices	5
No. layers	2
No. CIAs between HMMs	15
max in-degree	2
No. Obs	5 p.t.
No. Latent	3 p.t.

The results of the experiment is summarised by Figure 5, which shows the relative entropy to the generative ground-truth DIN over the number of training samples used.

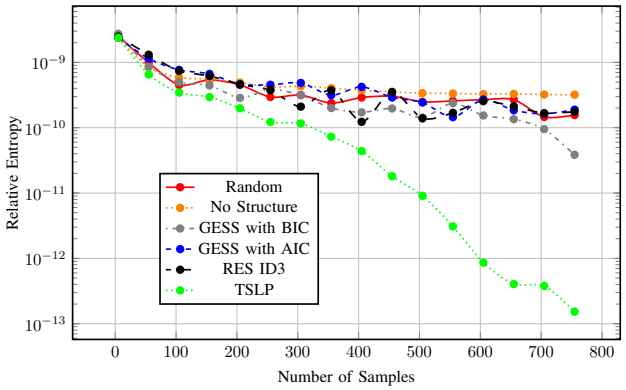


Fig. 5. The performance of parameter and structure learning tasks.

The results are averaged over 10 trials for various structure learning tasks:

- Random, which used a randomly generated structure for a DIN and learned the missing and observable parameters;
- No structure, which modelled each HMM as mutually independent to others and learned parameters;
- d-BIC with GESS, which used the d-BIC score with GESS and learned missing and observable parameters;
- d-AIC with GESS;
- RES ID3, which used the likelihood score but restricted recovered structures with an in-degree greater than 3 and also learned parameters;
- and finally, TSLP (Learning Procedure with the True structure), which used the ground-truth DIN structure, but relearned missing and observable parameters.

The parameters of the experiment were as follows: 20 EM iterations were used for learning and latent variables. There were 50 structure search iterations used to recover each model (5 random restarts when reached local optima, and used a tabu list of length 10). All learned variables used a Dirichlet Prior of 5. To allow for a manageable use of memory all DIN with over 5000 independent parameters were heavily penalised.

The reported results suggest that, on the one hand, when we have fewer samples we are better-off not using any structure, since fewer parameters allow us to generalise better. On the other hand, when we have a sufficient amount of data, then a random structure gives us more information than no structure at all. The reason the random structure does better is because the likelihood to the data of a structure with more conditional independence assumptions rather than fewer will also be greater.

The three penalty-based score methods do better than both random and no structure. However we find that the sensitivity of the d-BIC score to judge when to constrain the structure (based on the number of training sample) guides the selection of the independence assumptions and outperforms the d-AIC score and the restricting the in-degree method. As expected, knowing the true structure gives us the most information and thus outperforms all the methods as the number of observations increase.

V. CONCLUSION AND IMPLICATIONS

In this paper we empirically demonstrated a score-based structure learning procedure to learn a DIN to represent the influence relationships between partially observable stochastic processes.

Why we would want to learn a DIN depends on what it will be used for. On the one hand, if we are trying to identify the original DIN for knowledge discovery, then we will need to identify each of the original conditional independence assumptions of the ground-truth network. This means we will need to find the set $\mathcal{I}(\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}))$. This is not a pragmatic task since there are many perfect maps for $P^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}})$ that can be derived from $\mathcal{D}_{\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}}$.

Recognising $\mathcal{I}(\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}))$ from the set of structures from $\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}})$ will yield the same fit to the data. Therefore identifying the original ground-truth structure is not identifiable from $\mathcal{H}_{\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}}$. This is because the structures in the I-equivalent structure set all produce the same numeric likelihood (mutual information) for $\mathcal{H}_{\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}}$. Therefore, we should rather try to learn a set of structures that are I-equivalent to $\mathcal{G}^*$.

On the other hand, if instead we are trying to learn a DIN for density estimation (i.e. to draw probabilistic inferences), then we are interested in capturing the distribution $P^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}})$. If we can successfully construct such a distribution then we can reason about or sample new data instances.

There are two implications when learning a structure or density estimation: Firstly, Although capturing more independence assertions than specified in $\mathcal{I}(\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}))$ may still allow us to capture $P^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}$, our selection of more independence assumptions could result in *data fragmentation*. Secondly, selecting very sparse structures can restrict us to never being able to learn the true distribution $P^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}$ no matter how we change the parameters. However, often sparse DIN structures can be used to promote computational complexity savings [16].

ACKNOWLEDGEMENTS

This work is based on the research supported in part by the National Research Foundation of South Africa (Grant number: 121835).

REFERENCES

[1] R. Ajoodha, "Influence modelling and learning between dynamic bayesian networks using score-based structure learning," *The University of the Witwatersrand, Johannesburg (wirespace)*, 2019.

[2] R. Ajoodha and B. Rosman, "Learning the influence structure between partially observed stochastic processes using iot sensor data," in *Workshops at the Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[3] R. Ajoodha and B. Rosman, "Tracking influence between naïve bayes models using score-based structure learning," in *2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech)*. IEEE, 2017, pp. 122–127.

[4] R. Ajoodha and B. Rosman, "Learning the influence between partially observable processes using score-based structure learning," *Advances in Science, Technology and Engineering Systems Journal. Special Issue on Multidisciplinary Sciences and Engineering*, 2020.

[5] A. Anshel and D. A. Kipper, "The influence of group singing on trust and cooperation," *Journal of Music Therapy*, vol. 25, no. 3, pp. 145–155, 1988.

[6] R. Ajoodha, R. Klein, and B. Rosman, "Single-labelled music genre classification using content-based features," in *IEEE proceedings, Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech), 2015*, Nov 2015, pp. 66–71.

[7] R. Ajoodha, A. Jadhav, and S. Dukhan, "Forecasting learner attrition for student success at a south african university," in *In Conference of the South African Institute of Computer Scientists and Information Technologists 2020 (SAICSIT '20), September 14-16, 2020, Cape Town, South Africa. ACM, New York, NY, USA, 10 pages.* ACM, 2020.

[8] R. Ajoodha, S. Dukhan, and A. Jadhav, "Data-driven student support for academic successby developing student skill profiles," in *International Multidisciplinary Information Technology and Engineering Conference. ISBN: 978-1-7281-9519-9.* IEEE, 2020.

[9] T. Abed, R. Ajoodha, and A. Jadhav, "A prediction model to improve student placement at a south african higher education institution," in *2020 International SAUPEC/RobMech/PRASA Conference*. IEEE, 2020, pp. 1–6.

[10] J. Hatfield, G. J. Faunce, and R. Job, "Avoiding confusion surrounding the phrase "correlation does not imply causation"," *Teaching of Psychology*, vol. 33, no. 1, pp. 49–51, 2006.

[11] R. Opgen-Rhein and K. Strimmer, "From correlation to causation networks: a simple approximate learning algorithm and its application to high-dimensional plant gene expression data," *BMC systems biology*, vol. 1, no. 1, p. 37, 2007.

[12] T. Grinthal and N. Berkeley Heights, "Correlation vs. causation," *AMERICAN SCIENTIST*, vol. 103, no. 2, pp. 84–84, 2015.

[13] D. Commenges and A. Gégout-Petit, "A general dynamical statistical model with causal interpretation," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 71, no. 3, pp. 719–736, 2009.

[14] M. Bunge, *Causality and modern science*. Routledge, 2017.

[15] W. Salmon, "Scientific explanation and the causal structure of the world," 1984.

[16] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[17] D. Heckerman and D. Geiger, "Learning bayesian networks: a unification for discrete and gaussian domains," in *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 1995, pp. 274–284.

[18] A. Mohammadi and E. C. Wit, "Bayesian structure learning in sparse gaussian graphical models," *Bayesian Analysis*, vol. 10, no. 1, pp. 109–138, 2015.

[19] A. L. Madsen, F. Jensen, A. Salmerón, H. Langseth, and T. D. Nielsen, "A parallel algorithm for bayesian network structure learning from large data sets," *Knowledge-Based Systems*, vol. 117, pp. 46–55, 2017.

[20] X. Fan, C. Yuan, and B. M. Malone, "Tightening bounds for bayesian network structure learning." in *AAAI*, 2014, pp. 2439–2445.

[21] C. P. d. Campos and Q. Ji, "Efficient structure learning of bayesian networks using constraints," *Journal of Machine Learning Research*, vol. 12, no. Mar, pp. 663–689, 2011.

[22] S. Kok and P. Domingos, "Learning the structure of markov logic networks," in *Proceedings of the 22nd international conference on Machine learning*. ACM, 2005, pp. 441–448.

[23] J. B. Tenenbaum, C. Kemp, T. L. Griffiths, and N. D. Goodman, "How to grow a mind: Statistics, structure, and abstraction," *science*, vol. 331, no. 6022, pp. 1279–1285, 2011.

[24] I. Tsamardinos, L. E. Brown, and C. F. Aliferis, "The max-min hill-climbing bayesian network structure learning algorithm," *Machine learning*, vol. 65, no. 1, pp. 31–78, 2006.

[25] S.-I. Lee, V. Ganapathi, and D. Koller, "Efficient structure learning of markov networks using l_1-regularization," in *Advances in neural Information processing systems*, 2007, pp. 817–824.

[26] D. M. Chickering, D. Geiger, and D. Heckerman, "Learning bayesian networks is np-hard," Technical Report MSR-TR-94-17, Microsoft Research, Tech. Rep., 1994.

[27] D. M. Chickering, "Learning bayesian networks is np-complete," in *Learning from data*. Springer, 1996, pp. 121–130.

[28] D. M. Chickering, D. Heckerman, and C. Meek, "Large-sample learning of bayesian networks is np-hard," *Journal of Machine Learning Research*, vol. 5, no. Oct, pp. 1287–1330, 2004.

[29] J. Suzuki, "An efficient bayesian network structure learning strategy," *New Generation Computing*, vol. 35, no. 1, pp. 105–124, 2017.

[30] G. Schwarz *et al.*, "Estimating the dimension of a model," *The annals of statistics*, vol. 6, no. 2, pp. 461–464, 1978.

[31] S. Chen and P. Gopalakrishnan, "Speaker, environment and channel change detection and clustering via the bayesian information criterion," in *Proc. darpa broadcast news transcription and understanding workshop*, vol. 8. Virginia, USA, 1998, pp. 127–132.

[32] Y. Tamura, T. Sato, M. Ooe, and M. Ishiguro, "A procedure for tidal analysis with a bayesian information criterion," *Geophysical Journal International*, vol. 104, no. 3, pp. 507–516, 1991.

[33] J. Rissanen, "Stochastic complexity," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 223–239, 1987.

[34] A. Barron, J. Rissanen, and B. Yu, "The minimum description length principle in coding and modeling," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2743–2760, 1998.

[35] D. Geiger, D. Heckerman, H. King, and C. Meek, "Stratified exponential families: graphical models and model selection," *Annals of statistics*, pp. 505–529, 2001.

[36] D. Rusakov and D. Geiger, "Asymptotic model selection for naive bayesian networks," *Journal of Machine Learning Research*, vol. 6, no. Jan, pp. 1–35, 2005.

[37] R. Settimi and J. Q. Smith, "On the geometry of bayesian graphical models with hidden variables," in *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 1998, pp. 472–479.

[38] M. Koivisto and K. Sood, "Exact bayesian structure discovery in bayesian networks," *Journal of Machine Learning Research*, vol. 5, no. May, pp. 549–573, 2004.

[39] A. P. Singh and A. W. Moore, "Finding optimal bayesian networks by dynamic programming," 2005.

[40] T. Silander and P. Myllymaki, "A simple approach for finding the globally optimal bayesian network structure," *arXiv preprint arXiv:1206.6875*, 2012.

[41] D. Chickering, D. Geiger, and D. Heckerman, "Learning bayesian networks: Search methods and experimental results," in *proceedings of fifth conference on artificial intelligence and statistics*, 1995, pp. 112–128.

[42] W. Buntine, "Theory refinement on bayesian networks," in *Proceedings of the Seventh conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann Publishers Inc., 1991, pp. 52–60.

[43] A. Moore and M. S. Lee, "Cached sufficient statistics for efficient machine learning with large datasets," *Journal of Artificial Intelligence Research*, vol. 8, no. 3, pp. 67–91, 1998.

[44] K. Deng and A. W. Moore, "Multiresolution instance-based learning," in *IJCAI*, vol. 95, 1995, pp. 1233–1239.

[45] A. W. Moore, "The anchors hierarchy: Using the triangle inequality to survive high dimensional data," in *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 2000, pp. 397–405.

[46] P. Komarek and A. W. Moore, "A dynamic adaptation of ad-trees for efficient machine learning on large data sets." in *ICML*, 2000, pp. 495–502.

[47] P. Indyk, "Nearest neighbors in high-dimensional spaces," 2004.

[48] K. P. Murphy, "Dynamic bayesian networks: representation, inference and learning," Ph.D. dissertation, University of California, Berkeley, 2002.
[49] L. Rabiner and B. Juang, "An introduction to hidden markov models," *ieee assp magazine*, vol. 3, no. 1, pp. 4–16, 1986.
[50] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

An ISM-Band Automated Irrigation System for Agriculture IoT

Micah Bogdanoff
Department of Electrical and Computer Engineering
California State University Fresno
Fresno, CA, U.S.A
mibadvanced@mail.fresnstate.edu

Shahab Tayeb
Department of Electrical and Computer Engineering
California State University Fresno
Fresno, CA, U.S.A
tayeb@csufresno.edu

***Abstract*—In this paper, an inexpensive and user-friendly agriculture automation system is proposed by networking a collection of sensors and actuators to sense the moisture content of the soil and control the water valves for multiple irrigation zones. Using freely available frequencies in the Industrial, Scientific, and Medical (ISM) bands, multiple sensors and actuators can be networked to communicate with one another without the need to pay for subscriptions to existing cellular networks. Each sensor and actuator connect to a central communication node which connects the end-user to the sensors and actuators through a cloud server. The proposed system can act as a base for large-scale smart agriculture deployments.**

***Index Terms*—Agriculture, IoT, ISM, Automation**

I. Introduction

This agriculture automation system is designed to ease the life of farmers that must often wake up late at night to irrigate their crops. It gives farmers the ability to remotely control and automate their irrigation on a fixed time schedule, or dependent on the moisture content of the soil.

The automated irrigation system is set up by networking multiple microcontrollers to communicate with each other via 2.4GHz transceivers. Each microcontroller communicates with the central node which synchronizes the data with a cloud server platform. The user interface connects to the cloud server, allowing the end-user to remotely interact with the physical sensors and actuators that are in the field.

In this paper, we propose a scalable interconnection of sensors and actuators communicating through ISM bands. Such a system is very cost-effective and does not depend on a cellular infrastructure, unlike many of the existing smart agricultural deployments.

II. Background

In recent years, the Internet of Things (IoT) has become an ever-expanding concept to describe interconnection of end user devices [1] [2] [3]. This concept now includes automation in the Agriculture sector, which has had an increased demand in recent years, particularly for large agribusiness [4]. Wireless Sensor Networks (WSN) are made to collect data from the surrounding environment without a wired connections. For farms, this is especially useful because it is not feasible to run wires through the field where cultivators may damage the wire. A wireless Sensor Network is especially useful for farmers to remotely receive information regarding soil moisture content, air humidity, well depth indicator, and other agriculture related environmental data [19].

Data from each sensor node can be transmitted wirelessly on the Industrial, Scientific and Medical (ISM) frequency bands without special licensing [5]. The transceivers used in this experiment operate on the 2.4 GHz wavelength which is available to use without significant restrictions [6]. Each sensor and actuator node can be connected using various network topologies, however, the two most common are Star network and Mesh network. Each network topology has its own advantages, but a star network topology has the advantage of being relatively less complex and more scalable than a mesh network topology, but with less redundancies in case of signal loss [7] [8] [9].

Various terrains and obstructions in the field can attenuate the signal and influence the likelihood of the data successfully transmitting. To avoid propagation loss, the link budget of the 2.4GHz transceiver must be calculated to avoid data loss while still maintaining the power limitations of a Low Power Wide Area Network (LPWAN) for a given bandwidth [10]. To maintain a long range network while consuming a small amount of energy, it is necessary to send a small amount of data in sparse time increments [11] Various environmental sensors do not produce a significant amount of data, hence the data can successfully be transmitted and received over a longer distance. To ensure that the transmitted data successfully reaches the receiver, a handshake between the transmitter and receiver can be setup to confirm receipt of data [12].

With so many devices interconnected, a network bottleneck can potentially develop while communicating with the cloud server. To overcome this issue, fog computing can be utilized to analyze the data on the edge layer of the network prior to synchronizing the data with the cloud server [13]. The sensors connected to the Universal Node communicate through the Serial Peripheral Interface (SPI) of the microcontroller which is compatible with various sensors.

978-1-7281-9615-2/20$31.00 ©2020 IEEE

III. DESIGN

A. Frequency Choice

The aim of this system design is to create a Low Power Wide Area Network (LPWAN) that enables the devices to communicate with each other over a long range while using minimal power. For this paper, 2.4GHz transceivers were chosen instead of their sub-GHz counterparts because there are no license fees, no duty cycle restrictions, is available globally and the parts are cheaper, therefore is more economical and scalable [14].

In urban areas, 2.4GHz can potentially become congested, however, since rural farms have few, if any devices communicating on 2.4GHz wavelength, cross-technology interference is not a concern, making it a viable wireless communication band for an Agriculture IoT system [15] [16] [17].

B. Network Architecture

Fig. 1 depicts the process flow diagram of the system, which consists of several layers. The first layer is comprised of the Universal Nodes which have sensors and actuators connected to them. Sensors data is sent from the Universal Nodes to the Central Node wirelessly, and in turn, actuator commands are relayed on the same channels from the Central Node back to each Universal Node to activate the corresponding relays for each valve.

The Central Node is connected to the internet through a standard Wi-Fi network router which allows the data from the Central Node to synchronize with the cloud server. The data on the cloud server can be accessed and altered through a web application which serves as the user interface to access the data from the sensors and to control the actuators connected to the Universal Nodes on the first layer.

The Universal Nodes are wirelessly connected to the Central Node in a star network topology using the nRF24L01+ transceiver modules. Each transceiver can support up to 125 channels, allowing the Central Node to communicate with up to 125 Universal Nodes.

The Central Node was made with an ESP8266 microcontroller, which was chosen for its built-in internet connectivity and low cost. Although it consumes more power than its Arduino counterparts, the Central Node will be connected to a power outlet near the Wi-Fi router, serving as the gateway to the internet.

A closer look at the Universal Node in Fig. 1 shows the internal components which are comprised of an Arduino Nano microcontroller, a 3.7v battery, voltage converter, charge controller, solar cell, relay, and 2.4GHz transceiver. The entire setup is enclosed in a waterproof case which has the sensor and actuator connection ports on the side of the device as seen in Fig. 2.

In this paper, both a soil moisture sensor, and temperature sensor were connected to the Universal Nodes. A capacitive moisture sensor was chosen for its resistance to corrosion over time, unlike other moisture sensors that have two conducting probes which corrode within a short period of time.

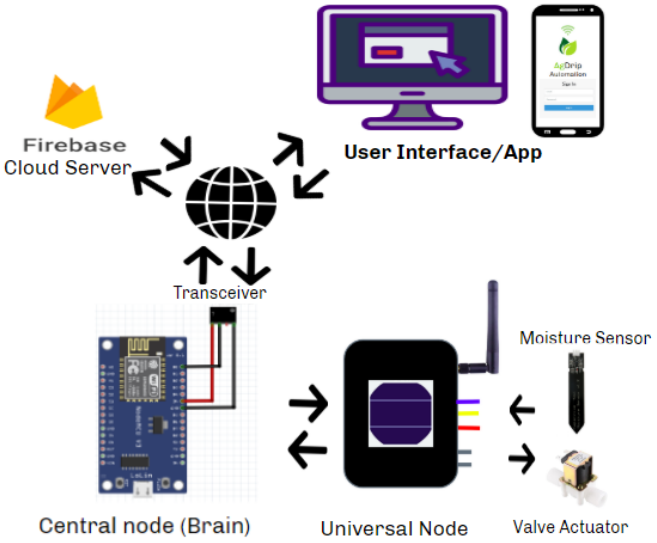


Fig. 1. Network Layout

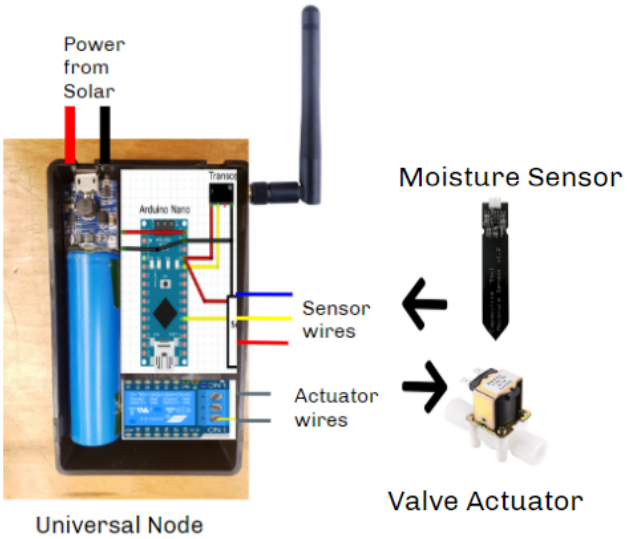


Fig. 2. Universal Node

To measure the outdoor temperature, a simple thermistor was connected to one of the Arduinos GPIO pins. A standard Negative Temperature Coefficient (NTC) thermistor decreases its resistance in response to an increase in room temperature. Using a simple voltage divider (1), the output voltage varies in relation to the resistance of the thermistor.

$$V_{out} = V_{in}\left(\frac{R2}{R1 + R2}\right) \tag{1}$$

With a thermistor (R1) in the $10k\Omega$ range, a $10k\Omega$ resistor was chosen for R2. The software was then calibrated to relate each voltage value to its corresponding temperature value.

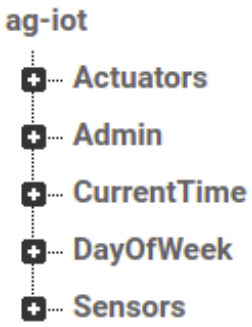


Fig. 3. JSON tree in Firebase

Fig. 4. Login Page

Data received by the Central Node is immediately synchronized with the cloud server Firebase. All the data values are stored in a JSON tree format on Firebase where each sensor value can be viewed, and the actuator values can be altered to control the valves and electrical relays. In Fig 3, a real-time database was set up in Firebase to store and synchronize the data so it can be accessed by the user interface.

The following login page from Fig. 4 is presented when logging into the user interface.

Password protection was implemented by setting parameters for the login entry to match the credentials from Fig. 5 in Firebase.

The main page of the user interface from Fig. 6 displays the real-time status of each sensors and actuator.

The values for each sensor are stored in the JSON tree format on Firebase as seen in Fig. 10.

To keep cloud server costs to a minimum, and to avoid a network bottleneck for future scalability, it is important to reduce the amount of data sent to the cloud server, and to minimize the amount of data processing on the cloud.

This is done by performing most of the computing at the edge layer by the Arduino Nanos. The thermistor voltage

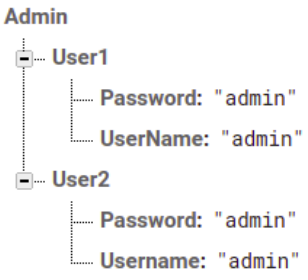


Fig. 5. Login branch expanded

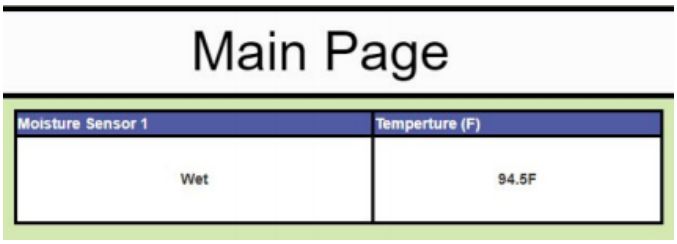


Fig. 6. Main Page

values are immediately converted to relevant Fahrenheit values before getting transmitted to the Central Node. Similarly, the moisture, flow indicator, and well depth values are all calculated prior to transmitting them to the Central Node which relays the information to the cloud server to be stored.

The subroutines were written to also provide compatibility with other analog and digital sensors. Synchronizing the cloud server with the Central Node allows for data to be altered to control the actuators. In the following expanded view of the actuator branch from Fig. 8, each actuator value is either a logic 0 or a logic 1. These values are controlled by connecting the cloud server to a user interface. Each actuator is turned on or off by activating a low voltage relay that is connected to a solenoid valve or to the primary pump.

To control the actuators, the user must be able to easily

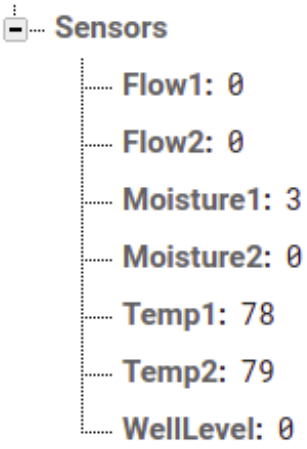


Fig. 7. Sensor branch expanded

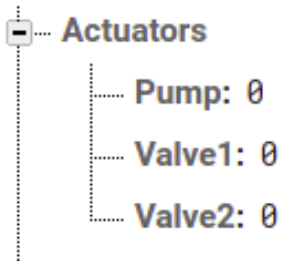


Fig. 8. Actuator branch expanded

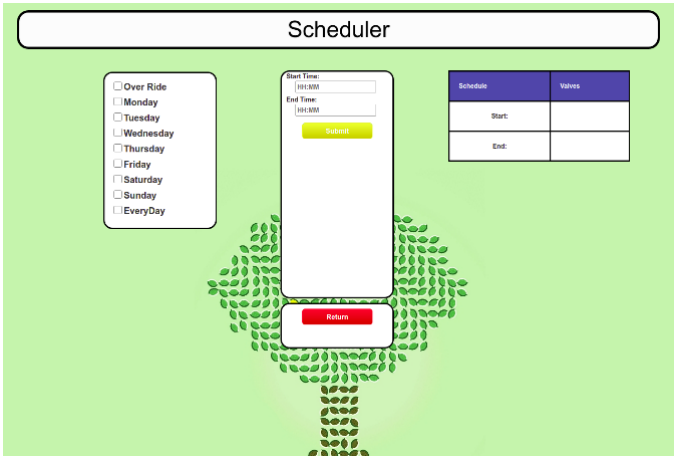


Fig. 9. Scheduler page

interact with the data through a user interface. The user interface JavaScript code was hosted locally on the computer while running the HTML web page from the google chrome browser. In the following image on Fig. 9, the scheduler and control interface are shown.

Each schedule value can be selected and modified on the webpage, which causes the cloud server to automatically update the scheduler data, as seen in Fig. 10.

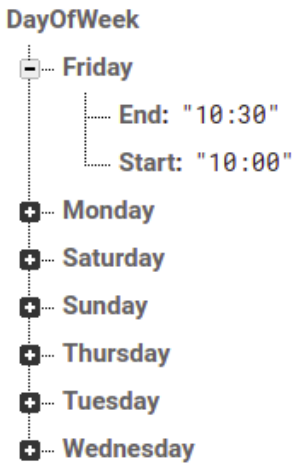


Fig. 10. Scheduler branch expanded

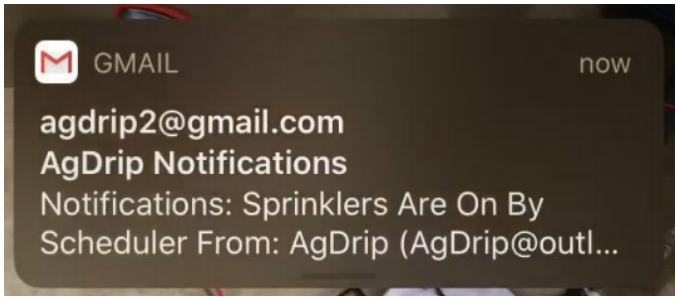


Fig. 11. mail Notification

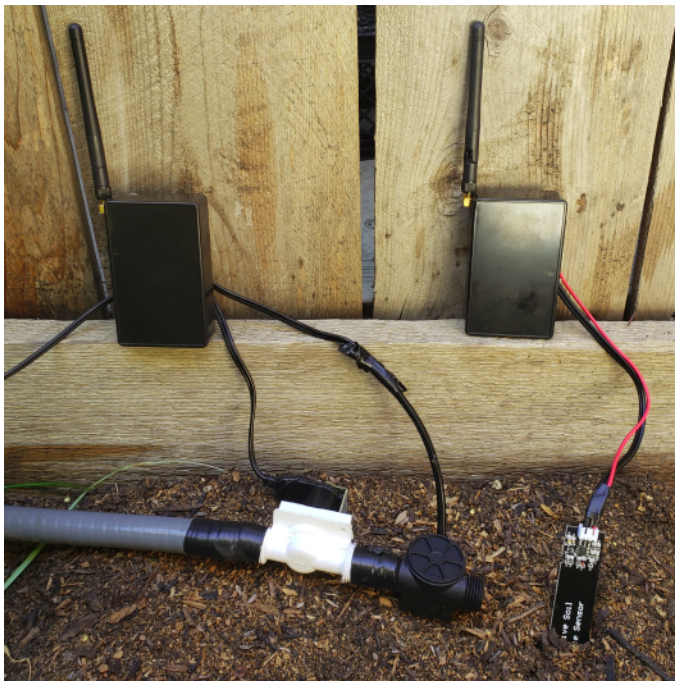


Fig. 12. Physical Setup

A notification system was also made to notify the user when the valves are activated. The notifications are sent via email, or can also be configured to send a text or a push notification, depending on the user's preference.

IV. Physical Setup

Fig. 12 shows the automated irrigation system prototype which was built and tested on a small garden as a proof of concept. The moisture sensor is connected to the Universal Node on the right, which relays the data back to the Central Node. On the left, the second Universal Node also sends the flow sensor data back to the Central Node. When a command is given to turn on the valve, the left Universal Node activates the relay which turns on the valve.

A short distance away, the Central Node receives the data from each sensor, and sends commands to the actuators. In Fig. 13, the Central Node, also enclosed in a case, connects to the internet via the NETGEAR router on the left.

Fig. 13. Central Node setup

TABLE I
FCC 2.4GHz Regulations (Point-to-Multipoint

Maximum Power (mW)	Maximum Gain (dBi)	EIRP (dBm)
1000	6	36
500	9	36
250	12	36
125	15	36
63	18	36
32	21	36
16	24	36

V. Regulations and Standards

A. RF Regulations

In the US, all radio transmission is regulated by the Federal Communications Commission (FCC). Some frequencies are not licensed, rather, they are available for anyone to use, so long as they adhere to the FCC regulations corresponding with that frequency.

The transceivers used in this paper operate in the ISM band of 2.4 GHz. The following table shows the maximum permissible radiated power depending on the gain of the antenna.

In addition to ensuring that the transceivers complied with FCC regulations, the hardware was confirmed to meet the regulations and standards set forth by the IEEE 802.11 and 802.15 standards which relate to WiFi and Wide Area Networks (WAN).

B. Intellectual Property

The Arduino Nanos used in this paper are under a General Public License (GPL), meaning it is required that any source code developed on an Arduino must be publicly available [18]. To comply with these regulations, a copy of the Arduino code was uploaded to GitHub where it is accessible to the public.

VI. Conclusion

Agriculture automation has become much simpler to manage by integrating various technologies. The ability to remotely gather real-time data from the fields, and to remotely activate irrigation has revolutionized agribusiness, increasing overall efficiency. Furthermore, the utilization of inexpensive transceiver modules operating on the free ISM bands reduces the overall cost to extend the internet of things into the agriculture sector. The results from this paper prove the concept of a feasible and easy to implement agriculture automation system which allows farmers to enjoy the benefits the remote accessibility and automation. With the same scalable system developed here, additional sensors and actuators can be added to reap the benefits that IoT has to offer.

References

[1] P. Pico-Valencia, J. A. Holgado-Terriza and X. Quiñónez-Ku, "A Brief Survey of the Main Internet-Based Approaches. An Outlook from the Internet of Things Perspective," 2020 3rd International Conference on Information and Computer Technologies (ICICT), San Jose, CA, USA, 2020, pp. 536-542, doi: 10.1109/ICICT50521.2020.00091.

[2] S. Tayeb, S. Latifi and Y. Kim, "A survey on IoT communication and computation frameworks: An industrial perspective," 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, 2017, pp. 1-6. doi: 10.1109/CCWC.2017.7868354

[3] S. Tayeb, Raste, N., Pirouz, M. and Latifi, S. "A Cognitive Framework to Secure Smart Cities," In MATEC Web of Conferences , 2018, (Vol. 208, p. 1).

[4] V. Grimblatt, "IoT for Agribusiness: An overview," 2020 IEEE 11th Latin American Symposium on Circuits & Systems (LASCAS), San Jose, Costa Rica, 2020, pp. 1-4, doi: 10.1109/LASCAS45839.2020.9068986.

[5] FCC Rules and Regulations, AIR802. [Online]. Available: https://www.air802.com/fcc-rules-and-regulations.html. [Accessed: 10-May-2020].

[6] Federal Communications Commission. [Online]. Available: https://transition.fcc.gov/Bureaus/EngineeringTechnology/Orders/1997/fcc97114.txt. [Accessed: 10-May-2020].

[7] H. Lee and K. Ke, "Monitoring of Large-Area IoT Sensors Using a LoRa Wireless Mesh Network System: Design and Evaluation," in IEEE Transactions on Instrumentation and Measurement, vol. 67, no. 9, pp. 2177-2187, Sept. 2018, doi: 10.1109/TIM.2018.2814082.

[8] O. Vondrouš, Z. Kocur, T. Hégr and O. Slavíček, "Performance evaluation of IoT mesh networking technology in ISM frequency band," 2016 17th International Conference on Mechatronics - Mechatronika (ME), Prague, 2016, pp. 1-8.

[9] H. Mamat, B. H. Ibrahim and M. P. Sulong, "Network Topology Comparison for Internet Communication and IoT Connectivity," 2019 IEEE Conference on Open Systems (ICOS), Pulau Pinang, Malaysia, 2019, pp. 1-5, doi: 10.1109/ICOS47562.2019.8975702.

[10] M. Najam ud din, "Indoor RF link budget analysis and design simulation of 2.4 GHz and 5.7 GHz dual band wireless transceiver,"2017 International Conference on Communication Technologies (ComTech), Rawalpindi, 2017, pp. 1-6, doi: 10.1109/COMTECH.2017.8065740

[11] A. Lavric and V. Popa, "A LoRaWAN: Long range wide area networks study," 2017 International Conference on Electromechanical and Power Systems (SIELMEN), Iasi, 2017, pp. 417-420, doi: 10.1109/SIELMEN.2017.8123360

[12] B. Babusiak, M. Smondrk and S. Borik, "Design of Ultra-Low-Energy Temperature and Humidity Sensor Based on nRF24 Wireless Technology," 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), Budapest, Hungary, 2019, pp. 397-401, doi: 10.1109/TSP.2019.8768890

[13] F. Jalali, S. Khodadustan, C. Gray, K. Hinton and F. Suits, "Greening IoT with Fog: A Survey," 2017 IEEE International Conference on Edge Computing (EDGE), Honolulu, HI, 2017, pp. 25-31, doi: 10.1109/IEEE.EDGE.2017.13.

[14] F. Wunsch, M. Ströer, M. Müller, H. Jäkel and F. K. Jondral, ”LPWAN Applications in the 2.4 GHz Band: A Viable Choice?,” 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), Chicago, IL, USA, 2018, pp. 1-5, doi: 10.1109/VTCFall.2018.8690804.
[15] T. Janssen, N. BniLam, M. Aernouts, R. Berkvens, and M. Weyn, “LoRa 2.4 GHz Communication Link and Range,” Sensors, vol. 20, no. 16, p. 4366, Aug. 2020.
[16] L. Polak and J. Milos, ”LTE and LoRa in the 2.4 GHz Band: Adjacent Channel Interference Issues,” 2020 30th International Conference Radioelektronika (RADIOELEKTRONIKA), Bratislava, Slovakia, 2020, pp. 1-4, doi: 10.1109/RADIOELEKTRONIKA49387.2020.9092382.
[17] A. Augustin, J. Yi, T. Clausen, and W. M. Townsley, ”A Study of LoRa: Long Range & Low Power Networks for the Internet of Things,” Sensors, vol. 16, no. 9s, pp. 1-18, Aug. 2020.
[18] M. Barbulescu et al., ”GNU GPL in studying programs from the Systems Engineering field,” 2011 RoEduNet International Conference 10th Edition: Networking in Education and Research, Iasi, 2011, pp. 1-4, doi: 10.1109/RoEduNet.2011.5993718.
[19] S. Tayeb, M. Mirnabibaboli, and S. Latifi. ”Cluster Head Energy Optimization in Wireless Sensor Networks.” Software Networking 2018.1 (2018): 137-162.

Perspective on the reliability behavior of intelligent transport systems during the transition phase from legacy vehicles to autonomous and connected ones: four-road intersections as a case study

Kamal AZGHIOU
EI Research Team
Université Mohammed Premier (UMP)
60000 Oujda, Morocco
k.azghiou@ump.ac.ma

Manal El MOUHIB
EDER Research Team
Université Mohammed Premier (UMP)
60000 Oujda, Morocco
m.elmouhib@ump.ac.ma

Abdelhamid BENALI
EI Research Team
Université Mohammed Premier (UMP)
60000 Oujda, Morocco
a.benali@ump.ac.ma

***Abstract*—We can find in the literature many research papers and technical reports tackling various problems related to the Autonomous Vehicles (AV) and Vehicle to everything (V2X) paradigms. However, safety guarantee of these technologies is the most critical problem to deal with because of its direct relation with human life. On one hand, we can find fewer works than the former, regarding the jointly study of the situations related to both AV and V2X. On the other hand, at our best of knowledge, there is no work investigating reliability in the transition phase when just a few of the Connected Autonomous Vehicle (CAV) are part of the traffic road. In this paper we explore this research gap, by analyzing and evaluating reliability of various CAV densities in the special case of four road intersection and for low traffic density. The results show that the convergence speed of reliability functions series to the reliability function of the system when only CAV populate the roads are not constant. This leads to the need of establishing solutions for the transitional phase to accompany the penetration process of the CAV in the automotive market.**

***Index Terms*—Autonomous Vehicles (AV), Vehicle to everything (V2X), Intelligent Transportation System (ITS), Reliability, Availability, Reliability Block Diagram (RBD), IoT**

I. Introduction

According to [1] road accidents are responsible for the death of 1.35 million person each year and the injury of 50 million ones around the world. Also, it's worth mentioning that 90% of these accidents are because of human errors [2]. Adding to that, the emissions of tens of millions of tonnes of CO_2 and the increase in global energy consumption are because of transportation. To resolve these issues, we need to assign driving tasks to the vehicle itself and/or maybe share it with several users. The autonomous driving paradigm can help to achieve the aforementioned objectives.

According to the Society of Automotive Engineers (SAE) classification, we can distinguish 6 levels of vehicle automation [3]. At the first level we do not have any kind of automation, the vehicle's driver is the only responsible of all driving tasks. The second level automates some of driving tasks such as speed regulation. Vehicles with automated manoeuvres such as parking assist and automated lane change capabilities belong to the level 3. while remaining vigilant, at the level four the vehicle driver can perform briefly other tasks than driving. At this level, the vehicle can autonomously perform driving tasks along longitudinal and lateral dimensions. At level 5, in some driving modes and under some conditions, the driver can turn away his attention to do other tasks as the vehicle moves. At the last level we achieve the ultimate vehicle automation where the driver do not assume any responsibility related to the driving tasks. Besides, the autonomous vehicle paradigm, we can find another one of no less of importance, namely the connected vehicle paradigm or V2X. In such paradigm, a connected vehicle can communicate with other connected vehicles, with the infrastructure (V2I), and with the Pedestrian (V2P) by the means of mobile telecommunication devices.

To achieve an almost perfect Intelligent Transportation System (ITS) both of the aforementioned paradigms need to cooper. Each of these paradigms fill the other's holes. As a critical system, ITS system safety attributes are the predominant concerns in ITS research areas [4], [5]. Reliability and latency of the network services are the most required Performance Key Indicators (PKIs) for a such type of mobile end terminal. All the Connected Autonomous Vehicles and the underlying infrastructure need to fulfil, among others, a pre-defined set of reliability requirements. However, assessing reliability in such an environment may face many difficulties such as systems heterogeneity, environmental constraints (Highway, Urban, Rural,...), weather behavior, stakeholders diversity to cite just a few.

We can find in the literature many works that deal with such problems (see section II). However, only few works tackle reliability issues of ITS systems in the transition stage that is characterized with the presence of vehicles belonging to different levels, and when V2I services are not available everywhere. This involves paying attention on how do reliability

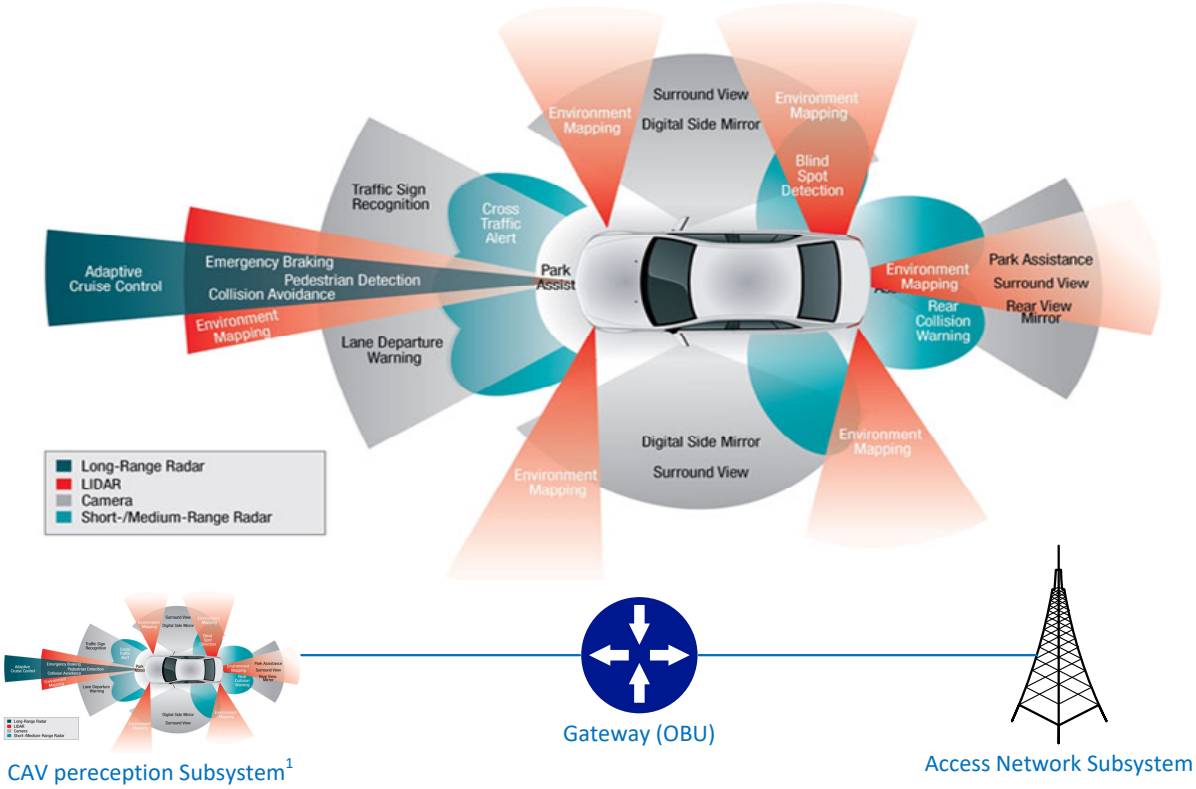
978-1-7281-9615-2/20/$31.00 ©2020 IEEE

behave in this time interval (that can be for many years) which depends on the penetration ratio of the CAV into the global automotive market. This paper aims to fill this gap by studying the reliability evolution of the CAV in presence of the legacy vehicles. We propose to investigate remarkable scenarios when the market penetration of the CAV is, respectively, 25%, 50%, 75% and 100%.

We organize this paper as follows : We propose in the section II a literature review of some related works. After that, in section III, and throughout, the considered system (subsection III-A), Reliability Modeling of the Autonomous Vehicles' perception subsystem (subsection III-B), Reliability Modeling of the connected Vehicles' embedded Gateway (subsection III-C), and the Reliability Modeling of the V2I Access Network (subsection III-D) subsections we present RBD-based modelling of the different subsystems. We dedicate the section IV to setup a simulation scenarios. Finally, results and discussions are the subject of the section V.

II. Related Work

The authors in [6], conducted a quantitative and a qualitative reliability analysis of the Basic Safety Messages (BSMs) in Vehicular Ad hoc Network (VANET). They introduced reliability model for safety applications (SA) using DSRC. As major results, they show that a safety application has higher reliability for small packets of sizes and failure to receive some BSM packets does not immediately affect the reliability of the safety application. Whereas, the authors in [7] explore the reliability of VANET using RBD to model the (On-Board Units) OBUs and the (Road Side Units) RSUs. Added to that, they modeled the availability and reliability of the communication channel using Continuous Time Markov Chains (CTMC) for Dedicated Short Range Communication (DSRC). The author in [8] considers the V2X scenario for the Ultra Reliable Low Latency Communication (URLLC) slicing in 5G. He proposes to improve both the Reliability and the Latency aspects by extending the resources slicing to the service and function slicing. To do so, he used tools like Euclidean Norm theory and Monte-Carlo simulations. More specifically, the authors in [9] conducted a System-level reliability analysis of V2X. They considered the Intersection Collision Avoidance (ICA) as safety application. They used Reliability Block Diagram (RBD) to analyse reliability components of the considered system such as hardware, software, maneuver. They considered the communication channel reliability to have a linear behavior depending on distance. [10] investigates the Reliability of fully Autonomous Vehicles (AV) using statistical testing. The authors reveal in their report that for to be Reliable, an AV must run hundreds of millions or billions of miles to justify



1 : https://www.machinedesign.com/mechanical-motion-systems/article/21836344/saved-by-the-sensor-vehicle-awareness-in-the-selfdriving-age

Fig. 1. Referral system used as basic modelling entity

its reliability. It is quality to notice that this report considered only AV with no V2X capabilities.

We can see that almost of these works focus only on the networking aspect of V2X systems reliability. Further, we observe that there is a need of works that consider CAV system reliability as a function of both its perception subsystem and its V2X environment. In addition, at the best of our knowledge there is no work assessing reliability of CAV system dealing with the transition phase of deployment when only a fraction of CAV are operational and the remaining vehicles are legacy ones. In this paper, we propose to model reliability of CAV systems by studying both its perception subsystem (AV aspect) and the networking subsystem (V2X aspect). After that, we conduct a simulation of 4 scenarios in which we investigate the impact of some remarkable CAV densities on the behavior of the overall reliability and thus on the road safety.

III. System Model

A. The considered system

The Fig.1 depicts the basic system which we will consider as a system of reference. We can recognize three main subsystems. The first one is the perception subsystem of CAV that consists of several components on which the CAV depends on when exploring its surrounding environment. The second one is the OBU, which represents the networking capability of the CAV. The third subsystem is the Access Network that can implement any technology supporting CAV applications such as DSRC and C-V2X. We mention that the Cloud subsystem is beyond the scope of this work.

In the remaining of the paper, we will adopt the Reliability Block Diagram (RBD) based framework of the Internet of Thing (IoT) and the reliability background theory detailed in [11]. We propose to model our referral system as three subsystems in series patterns as Fig.2 shows.

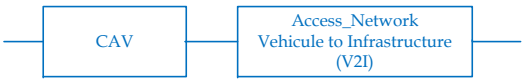


Fig. 2. Reliability Block diagram model for the global system : CAV block include the perception, and the OBU subsystems

B. Reliability Modeling of the the Autonomous Vehicles' perception subsystem

The perception subsystem of CAV comprises various sensors that collect environment related data. These sensors cover the vehicle throughout 360° around. We will use the model in Fig.1 where a CAV have one camera in each direction, radars in the front and the rear of the vehicle, LIDAR in the four directions, ultrasonic sensors in the front and the rear of the vehicle to detect the near objects, and finally a Long-Range Radar in the vehicle's front. The equation (1) assesses the reliability of the considered system:

$$\begin{aligned} R_{percep}(t) = (1 - R_{ultrasound}(t)) \times [1 - R^2_{radar}(t) \\ \times (1 - R_{radar}(t))^2 \times (1 - R_{lidar}(t))] \\ \times (1 - R_{camera}(t)) \times [(1 - R_{lidar}(t) \\ \times (1 - R_{lidar}(t)) \times (1 - R_{radar}(t))) \\ \times (1 - R_{camera}(t))]^2 \times (1 - R_{ultrasound}(t)) \\ \times (1 - R_{radar}(t)) \times (1 - R_{lidar}(t)) \\ \times (1 - R_{camera}(t)) \times (R_{wheelencoder}(t)) \end{aligned} \quad (1)$$

Where $R_x(t)$ is the reliability of the component x.

Depending on the situation, the CAV may require one or more sensors to perform a given driving task. We propose in our model to divide the proposed model of CAV into four subsystems, one for each direction of the vehicle (Front, rear, Front Left, Front Right). Given that, we model each of these subsystems by an RBD that reflects its structure. We consider that the sensors belonging to the same direction to be in a parallel pattern. The reason behind that is, in most of the time, these sensors cooper to perform the required driving task (Sensors fusion). The Fig.3 reveals the RBD model of the perception subsystem of the considered CAV.

C. Reliability Modeling of the connected Vehicles' embedded Gateway

CAV uses the OBU to communicate with other CAV (V2V), with the road infrastructure (V2I) or with the pedestrians (V2P). We propose to model the gateway part by the simplified RBD as shown in the Fig.4. We can observe that all the hardware part of the system is in series pattern.

The equation (2) assesses the reliability of the considered system:

$$\begin{aligned} R_{Gateway}(t) = R_{HMI}(t) \times R_{RCP}(t) \times R_{power}(t) \\ \times R_{GPS}(t) \times R_{DSRC}(t) \times R_{antenna}(t) \end{aligned} \quad (2)$$

Where $R_x(t)$ is the reliability of the component x.

D. Reliability Modeling of the V2I Access Network

We propose to model the reliability of the most interesting components in V2I access network subsystem, using the RBD paradigm. As the Fig.5 shows, the simplified model contains five blocks in series pattern. Some of these Block are the same as those in the OBU. We can recognize, DSRC antenna, the DSRC unit, the GPS unit in the RSU side, the power supply and the RSU router to communicate with the network infrastructure. We can observe that if any of the components fails all the subsystem fail.

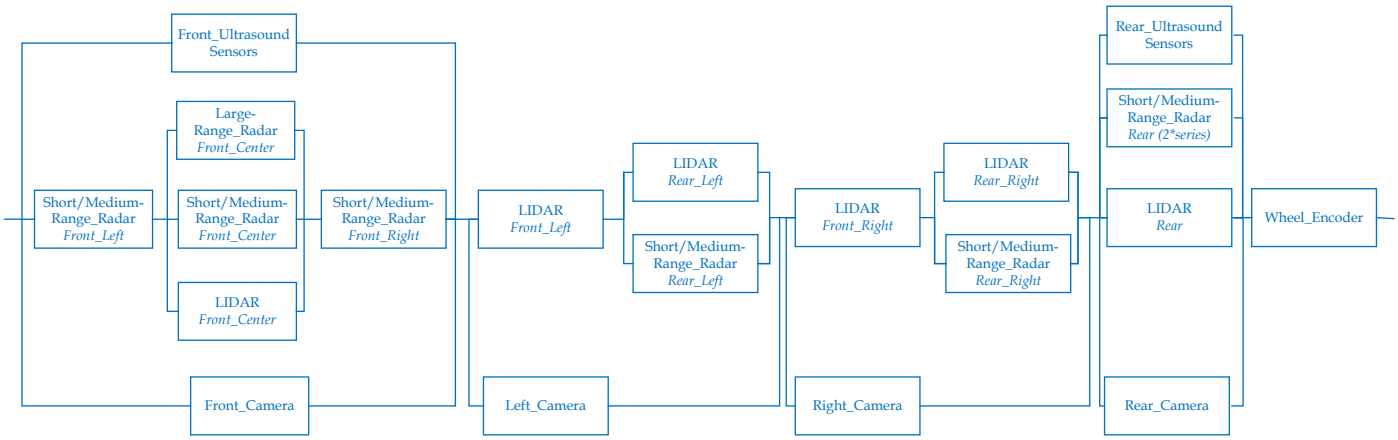


Fig. 3. The RBD model of the perception subsystem of the CAV

Fig. 4. The RBD model of the Gateway subsystem of the CAV

Fig. 5. The RBD model of the RSU of the CAV

The equation (3) assesses the reliability of the considered system:

$$R_{V2I}(t) = R_{antenna}(t) \times R_{DSRC}(t) \times R_{RSU}(t) \times R_{GPS}(t) \times R_{power}(t) \times R_{router}(t) \quad (3)$$

Where $R_x(t)$ is the reliability of the component x.

IV. Simulation Setup

A. Adopted scenarios and assumptions

To study the impact of CAV densities on improving the overall reliability of the CAV system, we propose to study four scenarios. We assume that all scenarios occur in urban intersections with four roads. Having a CAV and three non-CAV at the intersection, each on a separate road is more likely to happen if we have 25% of CAV density in the road traffic. We mention that we ignore the intersection scenarios when no CAV is engaged. Likewise, we can adopt the same approximation for 50 %, 75% and 100% CAV portions respectively and say that they correspond to two, three and four CAVs each crossing the intersection from a different road. However, is worth mentioning; in this context we do not consider the reliability analysis of communication channels for the sake of simplicity and because they are widely studied in other works. In addition, all the scenarios envisaged are assumed to involve low density traffic, which involves ignoring back end vehicles.

Fig.6 describes the scenario of 50% of density : :

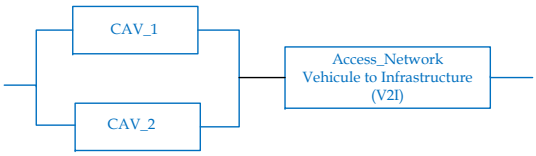


Fig. 6. The RBD model of the 50% of CAV density

Fig.7 describes the scenario of 75% of density :

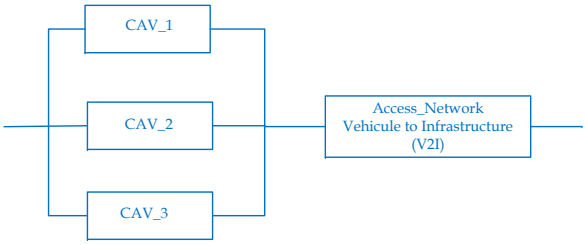


Fig. 7. The RBD model of the 75% of CAV density

Fig.8 describes the scenario of 100% of density

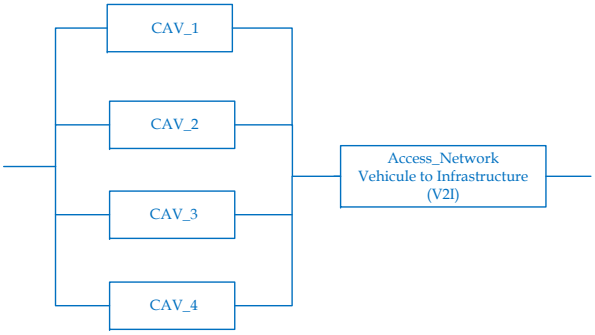


Fig. 8. The RBD model of the 100% of CAV density

To simulate the proposed RBDs, we utilise Reliability Workbench by injecting the failure rates specified in table (Table I) into models of the different blocks. We should mention that we choose the exponential distribution as failure model for all the blocks. We set mission time for the simulation to be 87600 hours (approximatively 10 years) which is reasonable duration to perceive changes in the automotive market. Also, we consider that all the components of the system are non repairable and independent one from each other.

V. Results and Discussions

Table II shows the got results after running the simulation in Reliability Workbench. we assessed reliability for the different aforementioned CAV densities in intersection situation as we mentioned previously. We consider the results for the following metrics :unreliability, unavailability, Mean Time to Failure (MTTF), total down time and Mean unavailability.

TABLE I
Simulation Data collected from [12], [13]

Component	Failure Rate
LIDAR	$1,8.10^{-6}$
Radar	$1,6.10^{-6}$
Ultrasound	$9,6.10^{-7}$
Camera	$2,9.10^{-6}$
Wheel Encoder	$4,4.10^{-7}$
HMI	$1,4.10^{-6}$
RCP	$2,8.10^{-6}$
Power Supply	$4,7.10^{-6}$
GPS	$1,4.10^{-6}$
DSRC Unit	$1,05.10^{-6}$
Antenna	$4,8.10^{-7}$
Router	$1,7.10^{-6}$

We can observe that unreliability and unavailability are the same for all CAV densities. This is because we considered system as non repairable. Also, it is clear from the table II that if CAV density increases the unreliability of the system decreases. The MTTF of the system makes a remarkable jump in the transition between 25% and 50% of CAV densities. Finally, we can notice that the MTTF is approximating 10 years.

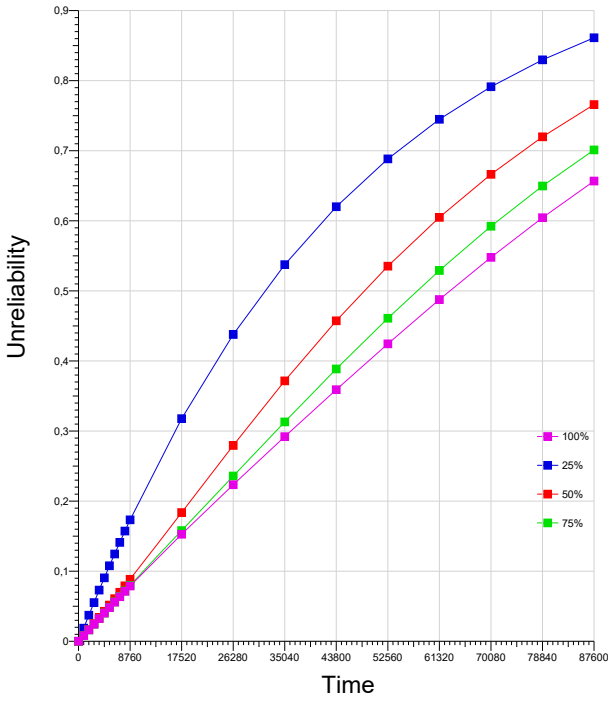


Fig. 9. Reliability behaviour in respect to the different CAV densities

In the Fig.9 we plot reliability function for each of the CAV density. The plot shows that for the considered road intersection case, there is a considerable gap between 25% density and the remaining ones in the time window [26280 hr,78840 hr]. In addition, we can see that, for the same time

TABLE II
SIMULATION RESULTS

Metric	density of 25%	density of 50%	density of 75%	density of 100%
Unreliability	0,86122	0.766211	0.70118	0.65662
Unavailability	0,86122	0.766211	0.70115	0.65667
MTTF	44389,2019	60282.5874	68718.5692	74047.4303
Total down time	48829,8549	37601.5328	32912.0223	30636.8559
Mean unavailability	0,55741	0.429241242	0.375708017	0.349735798

window as the unreliability of the overall system decreases and the CAV density increases the reliability gap between the system decreases.

These results reveal that a significant work need to be done in the transition phase to ensure road safety and support the CAV to penetrate the automotive market. When the market is in the stage of 25% of CAV density, we can think to increase reliability of the global system by taking some measures such as separating the CAV traffic from the legacy one or endowing the Infrastructure with necessary equipments to become a proxy for the CAV traffic.

A. Conclusion and further works

In this paper, we considered the study of reliability of a hybrid system composed by AV and V2X in the case of DSRC technology. We quantified reliability of the perception subsystem of the AV, the OBU and for the RSU. We defined the resulting system of putting all of those in series pattern as the referral system upon which we built our interpretation.

In the simulation part, we studied the effect of the traffic heterogeneity on the reliability behaviour of the global system. For that, we have made some simplifying assumptions, and we considered the special case of four road intersection in which several vehicles are attempting to cross it. We characterized the vehicles in the intersection to be CAV and non-CAV (or legacy vehicle). For more simplification, we considered just some remarkable density values (25%, 50%, 75% and 100%).

The results show that cooperative driving has a direct impact on the reliability of the overall system in the transition phase. More the density of CAV on the road traffic, more driving is reliable. Thus, we need to think about clever solutions to deploy in a such phase to increase the reliability of the global system.

We explored in this work a special case of road intersection. We can think to extend this investigation by considering more special situations like in highway, roundabout, and under different weather conditions.

REFERENCES

[1] "Global status report on road safety 2018," Tech. Rep. ISBN 978-92-4-156568-4, World Health Organization, Geneva, 2018.

[2] A. Haghi, D. Ketabi, M. Ghanbari, and H. Rajabi, "Assessment of Human Errors in Driving Accidents; Analysis of the Causes Based on Aberrant Behaviors," *Life Science Journal*, vol. 11, May 2014.

[3] On-Road Automated Driving (ORAD) committee, "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles," tech. rep., SAE International, 2018.

[4] L. Ye and T. Yamamoto, "Modeling connected and autonomous vehicles in heterogeneous traffic flow," *Physica A: Statistical Mechanics and its Applications*, vol. 490, pp. 269–277, 2018.

[5] S. Feng, Y. Feng, X. Yan, S. Shen, S. Xu, and H. X. Liu, "Safety assessment of highly automated driving systems in test tracks: A new framework," *Accident Analysis & Prevention*, vol. 144, p. 105664, 2020.

[6] A. Dabboussi, R. Kouta, J. Gaber, M. Wack, B. El Hassan, and L. Nachabeh, "Dependability overview for autonomous vehicles and reliability analysis for basic safety messages," in *2018 Sixth International Conference on Digital Information, Networking, and Wireless Communications (DINWC)*, (Beirut), pp. 86–91, IEEE, Apr. 2018.

[7] S. Dharmaraja, R. Vinayak, and K. S. Trivedi, "Reliability and survivability of vehicular ad hoc networks: An analytical approach," *Reliability Engineering & System Safety*, vol. 153, pp. 28–38, Sept. 2016.

[8] X. Ge, "Ultra-Reliable Low-Latency Communications in Autonomous Vehicular Networks," *IEEE Transactions on Vehicular Technology*, vol. 68, pp. 5005–5016, May 2019.

[9] Z. Wu, X. Zeng, and H. Yang, "System-Level Reliability Analysis of Cooperative Driving with V2X Communication for Intersection Collision Avoidance," *Transportation Research Record: Journal of the Transportation Research Board*, p. 036119812091975, May 2020.

[10] N. Kalra and S. M. Paddock, "Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?," p. 15.

[11] K. Azghiou, M. El Mouhib, M.-A. Koulali, and A. Benali, "An End-to-End Reliability Framework of the Internet of Things," *Sensors*, vol. 20, p. 2439, Apr. 2020.

[12] A. Dabboussi, *Dependability Approaches for Mobile Environment: Application on Connected Autonomous Vehicles.* PhD thesis.

[13] "Nonelectronic Parts Reliability Data 2016," Tech. Rep. NPRD-2016, Quanterion Solutions Incorporated, Utica, NY 13502-1311, 2016.

Design and Development of IOT Based Prototype Police Barricade

Antriksh Mangal
Department of Electrical Engg.
Delhi Technological University
New Delhi, India
antrikshmangal@gmail.com

Anirudh Goyal
Department of Electrical Engg.
Delhi Technological University
New Delhi, India
anirudhgoyal136@gmail.com

A.R.Kulkarni
Department of Electrical Engg.
Delhi Technological University
New Delhi, India
ashishkulkarni@dtu.ac.in

***Abstract*—With growing concerns over road traffic mismanagement and rampant untraced vehicle movement, it becomes imperative to deploy countermeasures. In this paper, we propose a solution for installation of contact-free road barricading system which employs motorized barricades and automatic vehicle e-pass authentication. To provide economical viability, the system is implemented using low cost hardware. Motorization is carried out using DC motor modulated by microcontroller and vehicle verification by computer vision techniques and scraping data from webpages of authorities governing road traffic.**

***Keywords*— ANPR, Haar Cascade, OCR, Arduino Uno, Raspberry Pi, Automatic Barricade, QR Code Reader**

I. Introduction

The current Covid-19 pandemic situation requires imposing restrictions on public commutes and limiting close physical interactions. Electronic passes for private vehicles emerge as a viable solution for managing road traffic. Present methods of implementation, however, require human to human contact for document verification which presents the counter-intuitiveness in the process. Proposing an alternate solution involving automation of the in-effect manually controlled barricade systems installed by traffic police.

When the commuter reaches the barricades, the camera placed 2 meters above the ground takes an image of the e-pass QR code. The camera placed at 50cms above the ground takes an image of the front license plate. The program on the Raspberry Pi first reads the QR code and stores the registration number, license plate number and confirms the pass validity. It then reads the license plate image and runs the Optical Character Recognition Algorithm [1] and compares it to the stored number from e-pass. After matchmaking, a trigger is sent to the Arduino Uno module to initiate the opening of barricade. Also, the driver's details along with the timestamps are recorded in an on-board database which can be uploaded on the main database at a later stage.

As the trigger is received by the Arduino module, it starts transmitting the control signals to the DC motor using the calculated torque and speed parameters. The barricade remains open till the vehicle passes. This is achieved with the help of an IR sensor placed on the barricade which detects vehicle movements.

II. Hardware Description

A. Barricade

Original design is meant for on-site installation on the traffic police barricades of dimensions 210cm x 150cm. However, due to the ongoing Covid-19 pandemic and henceforth, inaccessibility of the full-size barricades, we have designed a scaled-down model with $1/5^{th}$ proportions to the real equipment. The model barricade is designed from PVC pipes fitted in a rectangular shape of dimensions 42cm x 30cm. 2 such frames are joined with a gap of 5cm and 4 wheels of diameter 3cm are attached using 8cm long axles. The front axle is connected to the dc motor through a gear-train.

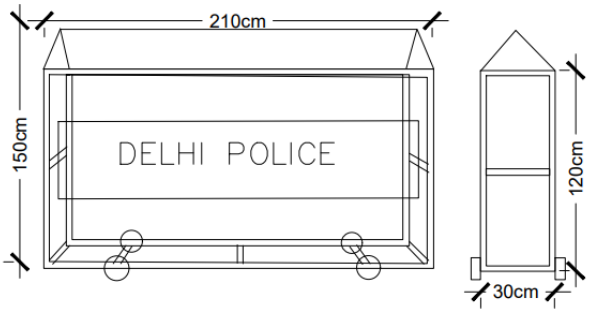


Fig.1 Delhi Police Barricade Dimensions

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Fig.2 Prototype Barricade Model

B. Arduino Uno

An open-source microcontroller board developed by Arduino.cc. It is based on the ATmega328P microchip, has a flash memory of 32KB, SRAM of 2KB and EEPROM of 1KB and works on a 16 MHz ceramic resonator (CSTCE16M0V53-R0). It has 14 digital IO pins, 6 analogue inputs, USB power jack, ICSP header, reset button and can be powered by a USB cable or an external 9V battery. Permissible DC Current of 20mA for IO pins and 50mA for 3.3V pin [2].

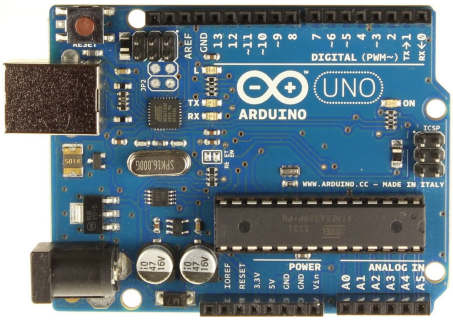


Fig.3 Arduino Uno Module
(Source: https://www.distrelec.biz/en/microcontroller-board-uno-arduino-a000066/p/11038919)

C. Raspberry Pi 3B+

Third in a series of single-board computers developed by the Raspberry Pi Foundation. It features a Broadcom BCM2837 SoC with a 1.2 GHz 64-bit quad-core ARM Cortex-A53 processor, 512 KB shared L2 cache, a dual-band IEEE 802.11b/g/n ac WiFi module, a Bluetooth 4.2 module and Gigabit Ethernet limited to 300 Mbps. It has a RAM of 1GB [3], [4].

Fig.4 Raspberry Pi Model 3B+
(Source: https://robu.in/product/raspberry-pi-3-model-b-bcm2837b0-soc-iot-poe-enabled/)

D. Camera

A USB video device based on CMOS technology, providing resolution of 1920x1080 pixels at frame-rate of 30fps. It is fit with a plastic lens that can be manually moved in and out to focus the camera.

Fig.5 OV2640 2Mp HD CMOS Camera
(Source: https://www.zigobot.ch/en/sensors/cam/ov2640-2mp-hd-cmos-camera-module-w-adapter-board-jpeg-out-detail.html)

E. DC Motor

A standard 130 type DC motor having an operating voltage range from 4.5V to 9V, rated voltage of 6V, maximum no-load current 70mA, a no-load speed of 3600 RPM, rated current of 250mA and rated torque of 10 gm-cm.

Since for this application a high torque dc motor is required, so to use the given motor, a gearbox with a gear ratio of 1:60 has been used. The gearbox increases the torque from 10 gm-cm to 500 gm-cm and consequently the speed reduces from 3600 RPM to 60 RPM.

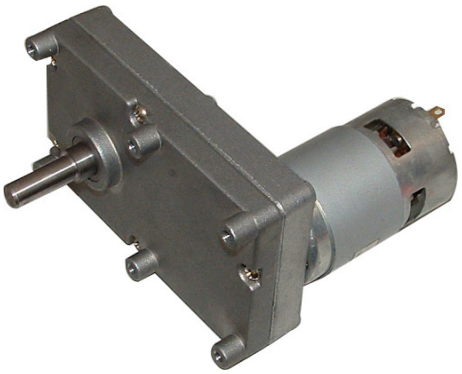


Fig.6 High torque geared DC Motor (Source: http://rcduniya.com/product/techtonics-12v-dc-square-gear–geared-motor-150-rpm-high-torque)

III. Design Description

A. QR Code Scanner

For detection and reading the QR code [5], [6] on e-pass, the python Pyzbar package is used on Raspberry Pi. After reading, a URL is produced which leads to an HTML webpage containing the e-pass verification number, validity time period, owner identity and vehicle registration number which are retrieved by the software using web-scraping tools.
Sample E-pass has been collected from Government of NCT of Delhi [7]. Some details in the image have been hidden to maintain privacy and integrity of data.

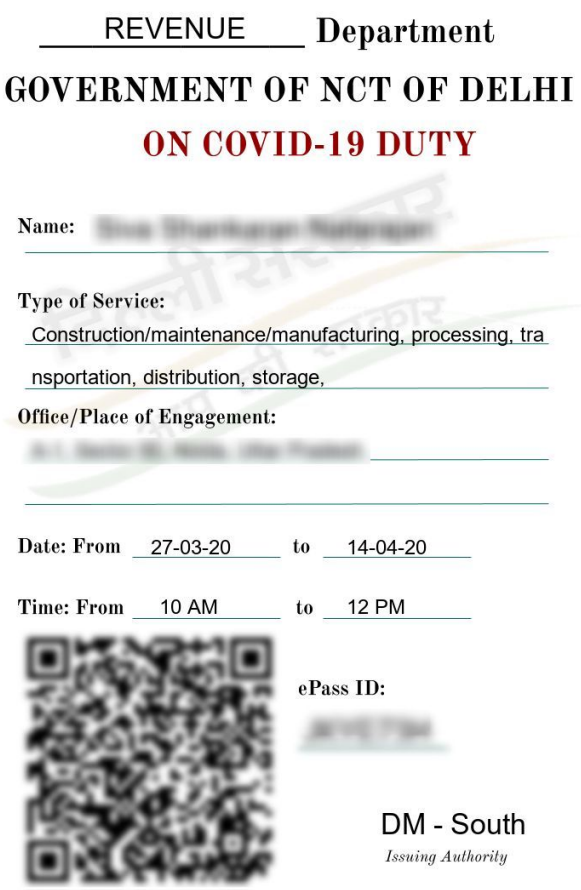

REVENUE Department
GOVERNMENT OF NCT OF DELHI
ON COVID-19 DUTY
Name:
Type of Service:
Construction/maintenance/manufacturing, processing, tra
nsportation, distribution, storage,
Office/Place of Engagement:
Date: From 27-03-20 to 14-04-20
Time: From 10 AM to 12 PM
ePass ID:
DM - South
Issuing Authority

Fig.7 Sample E-Pass

B. License Plate Recognition

For detecting the vehicle license plate, Haar Cascade feature detection algorithm [8], implemented using Python [9], is used which is fine-tuned for Indian license plate recognition. After generating the bounding box for the plate, a recurrent neural network OCR Algorithm [10], [11], [12] is run to read the characters on the cropped out license plate.

C. Raspberry Pi to Arduino Communication

After verifying the documents, a trigger is sent to the Arduino from Raspberry Pi [13]. For enabling wireless communication, we use the HC-05 Bluetooth module [14] for Arduino. The Pyserial library is used for sending triggers in the form of logic 0 (Rest state) and logic 1 (Open barricade) from Raspberry Pi.

Fig.8 HC-05 Bluetooth Module (Source: https://www.jsumo.com/hc-05-bluetooth-module-serial-transceiver-module)

D. Barricade Operation

The front axle of the barricade is fitted with the DC motor through a gear, which operates at 5V DC. The direction of motor operation is controlled using 2 relays which receive the control signal from the Arduino module. A rated torque of 10 gm-cm delivered by the DC motor allows the prototype barricade to operate at a linear speed of 0.1 m/s.

The opening/closing operation completes in 3 seconds covering a distance of 0.3 meters (equal to the length of the prototype barricade model). An IR sensor placed on the barricade detects if the vehicle has passed or not. On receiving the signal from the IR sensor, the closing process is initiated.

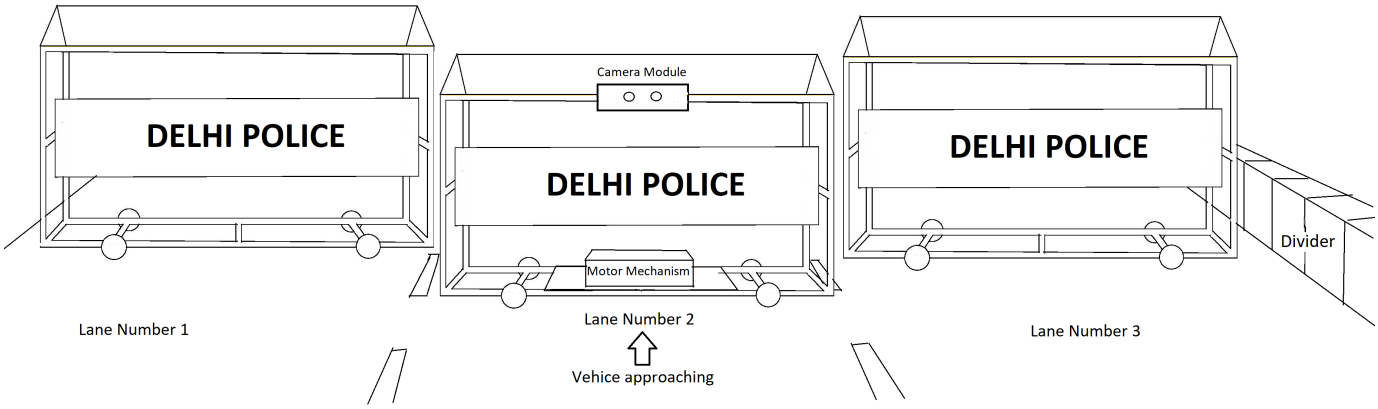


Fig.9 Barricades at standstill position

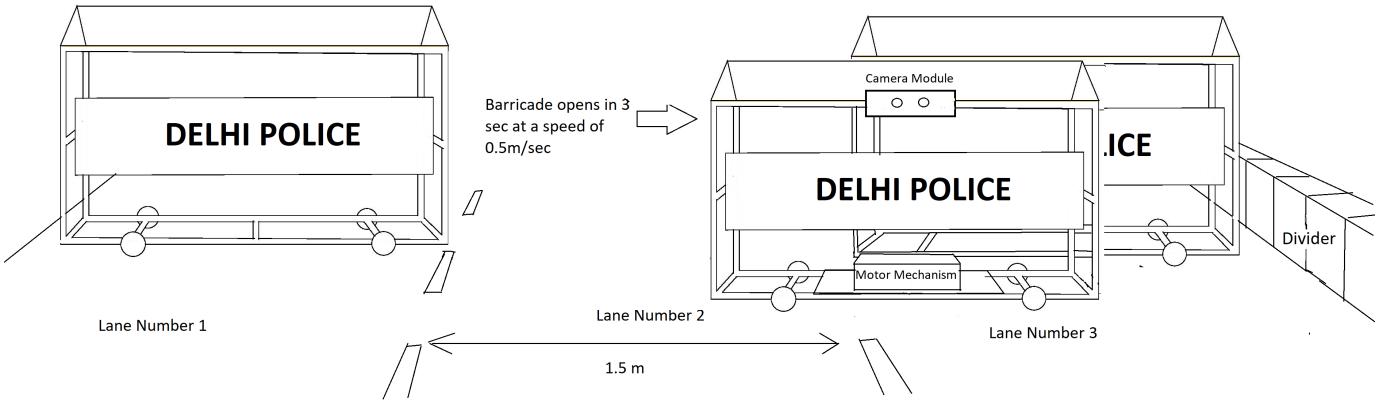


Fig.10 Barricades during opening operation

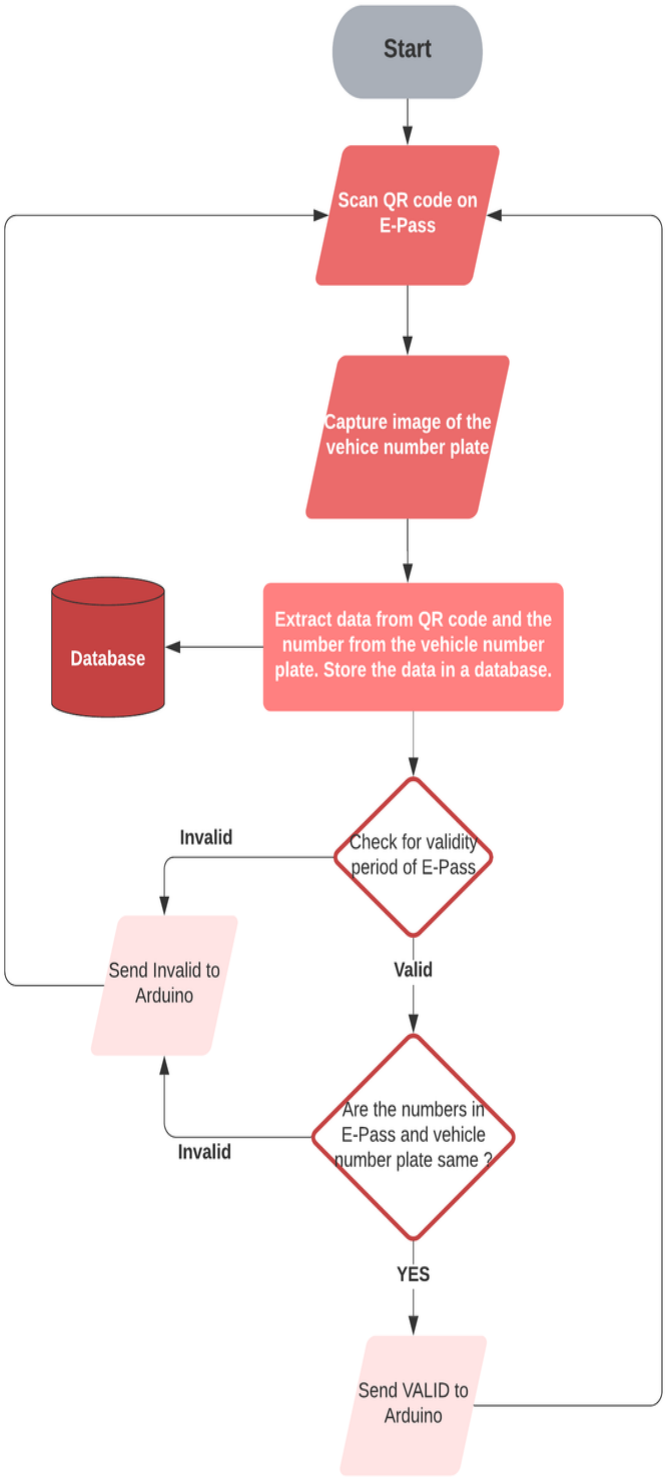


Fig.11 Raspberry Pi Operation Flowchart

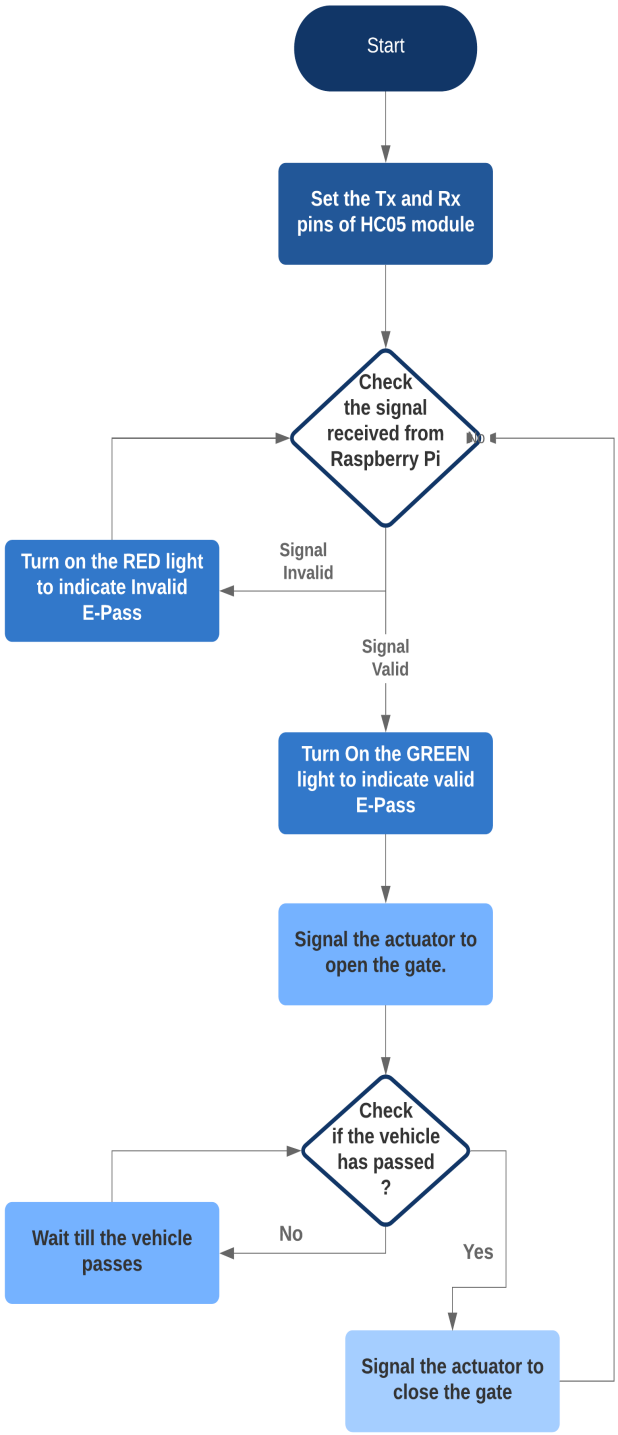


Fig.12 Arduino Operation Flowchart

IV. Conclusions

The automation of vehicle barricades and e-pass documents verification is implemented using Arduino microcontroller receiving trigger inputs from Raspberry Pi computer. Since no physical interaction is required for operating the setup, it adheres to the social distancing norms. Furthermore, the components are used to upgrade the existing systems and, hence, it is an economical solution.

V. Discussions

The OCR algorithm has a recurrent neural network at the core, and inherently has room for improvements to the accuracy metrics. It has a word-level accuracy of 70% [15]. Also, complex QR codes require well-focused image of e-pass document for decryption, thus revealing the speed-precision trade-off limitations.

The QR code and number plate recognition algorithm work on a continuous stream of live images even if no object is available for detection. A manual trigger is provided as an ultimate interrupt to circumvent any software malfunction situations, or to bypass the document verification process in case of exigencies.

The setup can be employed in vehicle-parking for automatic assignment of lots and defray through electronic payment methods. Also, home automation can be carried out by upgrading the mechanical gates to motorized ones.

Acknowledgments

We would like to thank the Head of Department and management for providing us with all facilities and support and special thanks to our guide for their encouragement and support.

References

[1] P. Kulkarni, A. Khatri, P. Banga, and K. Shah, "Automatic number plate recognition (anpr) system for indian conditions," *2009 19th International Conference Radioelektronika*, 2009.

[2] Arduino.cc, *Arduino - Uno - Rev3*, https://store.arduino.cc/usa/arduino-uno-rev3.

[3] raspberrypi.org, *Raspberry Pi 3 Model B+*, https://www.raspberrypi.org/products/raspberry-pi-3-model-b-plus/.

[4] Wikepedia, *Raspberry Pi*, https://en.wikipedia.org/wiki/Raspberry_Pi.

[5] K. Saranya, R. S. Reminaa, and S. Subhitsha, "Modern applications of qr-code for security," in *2016 IEEE International Conference on Engineering and Technology (ICETECH)*, 2016, pp. 173–177.

[6] M. E. Çoban, B. Çubukçu, R. Yayla, and U. Yüzgeç, "Raspberry pi based robot application using qr code: Qr-robot," in *2019 4th International Conference on Computer Science and Engineering (UBMK)*, 2019, pp. 119–123.

[7] G. of NCT of Delhi, *Sample E-Pass*, https://verify.epass.jantasamvad.org/epass/.

[8] Y. Zhao, J. Gu, C. Liu, S. Han, Y. Gao, and Q. Hu, "License plate location based on haar-like cascade classifiers and edges," in *2010 Second WRI Global Congress on Intelligent Systems*, vol. 3, 2010, pp. 102–105.

[9] G. Van Rossum and F. L. Drake, *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009.

[10] R. Smith, "An overview of the tesseract ocr engine," in *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, vol. 2, 2007, pp. 629–633.

[11] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[12] T. E. Oliphant, *A guide to NumPy*. Trelgol Publishing USA, 2006, vol. 1.

[13] V. Sandeep, K. L. Gopal, S. Naveen, A. Amudhan, and L. S. Kumar, "Globally accessible machine automation using raspberry pi based on internet of things," in *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2015, pp. 1144–1147.

[14] M. Wright, *HC-05 Bluetooth + Arduino: Includes the ZS-040*. CreateSpace Independent Publishing Platform, 2016, vol. 1.

[15] C. Patel, A. Patel, and D. Patel, "Optical character recognition by open source ocr tool tesseract: A case study," in *Int. J. Comput. Appl.*, vol. 55, no. 10, 2012, pp. 50–56.

A simple miniaturized wideband equilateral triangular microstrip patch antenna

Murari Shaw
Dept. of Electronics & communication Engg.
Institute of Engineering & Management-Salt Lake
Kolkata-700091, West Bengal (India)
zmz_shaw@yahoo.co.in

Malay Gangopadhyay
Dept. of Electronics & communication Engg.
Institute of Engineering & Management-Salt Lake
Kolkata-700091, West Bengal (India)
malay.ganguly@iemcal.com

Abstract*— A simple, very small equilateral triangular microstrip patch antenna (ETMPA) having sides 15mm with an I-slot of optimum length and position on the radiating patch has been designed to generate a wide frequency band. Impedance bandwidth generated by the antenna is 610MHz 10.8%, (5.34-5.95GHz) for ($S_{11} \leq -10$ dB) having broadside radiation pattern with peak gain of 1.66dB. The designed antenna covers Weather Radar application band(5.35-5.45)GHz, Radiolocation and Military application band (5.47-5.57) GHz, Maritime Radar application band (5.57-5.65)GHz, WLAN-Europe(5.470-5.725)GHz, Wi-Fi 5.8(5.725– 5.850)GHz and Dedicated Short-Range Communications(DSRC) band (5.850-5.925)GHz. FR4 material having relative permittivity(ε_r) 4.4 and loss tangent(tanδ) 0.02 has been used as dielectric substrate and thin copper sheet used as radiating patch and ground in the designed antenna. Prior to physical fabrication and measurement, the proposed antenna has been designed and analyzed using commercial High Frequency Structural Simulator (HFSS) ver13 software. Good matching has been observed between measured and simulated results.

Keywords—*miniaturization, wideband, resonant frequency.*

I. INTRODUCTION

From the inception of Microstrip Patch Antenna (MPA), huge research has been done in this area and now it has become the most usable type of antenna having wide area of application [1, 2].

With rapid growth in our modern communication technology, requirement of smaller and wide band (multiple bands applicable) MPA also increased and it becomes highly suitable to fulfill the demand of emerging communication area and technology because it has many attractive advantages like light weight, very small size, flexibility in shapes, cheaper cost and ease of fabrication [3-6]. With so many advantages, microstrip patch antenna also has some disadvantages like small bandwidth, low gain, lower efficiency and low power handling capacity [7-9].

The huge drawback of conventional MPA is its small bandwidth which for practical application, needs to be increased. Lot of research work has been performed in recent years to evolve techniques through which the bandwidth of MPA can be increased. Bandwidth-improvement techniques are- (i) Use of thick substrates having lower dielectric constant [10], (ii) Stacked patches [11], (iii) Addition of co-planar parasitic elements and (iv)Slotting of radiating patch.

Antenna designed with thick substrate having low dielectric constant can generate wide bandwidth but, due to increased height, length of probe feed will also increase, causing large probe inductance. In order to remove this large probe inductance, a capacitance can be incorporated by introducing a concentric annular cut around the probe feed [12], but this will make the design very complex and difficult to fabricate. Antenna designed using stacked patches can generate wide bandwidth but this design has drawbacks due to large size, complexity and difficulty in fabrication. Addition of co-planar parasitic elements along with radiating patch or slotting of radiating patch generates additional resonance that, together with the main resonance produces wide bandwidth response [13], but addition of co-planar parasitic elements with radiating patch increases the size and also makes the antenna more complicated to fabricate. In [14], an ETMPA having 50mm sides with two bent slots has been designed and by changing the location and lengths of two bent slots, the designed antenna has adjustable frequency ratio in the range of 1.2 to 1.43. Similarly, a dual frequency ETMPA having 42mm sides with three bent slots of 60^0 near each corner has been designed in [15], and by changing the lengths of three bent slots, the designed antenna has tunable frequency ratio varying in the range of 1.4 to 2. In [16] an ETMPA having 91mm sides with an optimum V-shaped slot have been constructed for dual operation at 1524 MHz and 2342 MHz frequencies. Another dual frequency ETMPA with a slit of variable length and position at the bottom side controlling frequency ratio which changes from 1.201 to 1.563 of the two operating frequencies of the antenna having 48mm sides have been designed in [17].Also, an equilateral triangular microstrip patch antenna excited using proximity fed method having 106mm each sides with bandwidth of 498MHz (37.53%) as well as peak gain more than 8dBi has been designed in [18], and In order to increase the bandwidth along with gain of radiating patch without changing its surface area, gap-coupled structure has been used with 30^0 sectoral patches in the designed antenna. In [19], a miniature stacked wideband MPA has been designed using Sierpinski fractal geometry method having enhanced bandwidth with optimum radiation efficiency and due to introduction of slot of triangular shape on both triangular active as well as parasitic patches, the resonant frequency of the antenna reduced. This slot introduction process repeated several time in order to miniaturize the traditional wideband stacked microstrip patch antenna. Similarly, a L-shaped probe fed wideband triangular microstrip patch antenna having 28mm sides and a thin foam layer to support radiating patch has been designed in [20] with 42% impedance bandwidth and 6dBi maximum gain.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

The growing communication technology causes miniaturization of communication equipments and devices. There is also commercial requirement of smaller antenna with wide frequency bandwidth that covers multiple application bands with low cost of fabrication. Therefore to meet the above requirement a very small equilateral triangular microstrip patch antenna has been proposed which has same radiation properties as rectangular microstrip antenna with less area [21-23]. Each side of the equilateral triangular radiating patch is 15mm and complete size of the antenna is (22x22x3.2) mm^3 that generates wide frequency band. The wide frequency band covers six application bands as well as very simple and cheaper in cost to fabricate. The radiating patch of the antenna has been loaded with an I-slot of optimum length and breadth to meet the design requirement. The designed antenna generates wideband of 610MHz 10.8% from 5.34GHz to 5.95 GHz having broadside radiation pattern throughout entire band width with peak gain of 1.66dB.

So far, no equilateral triangular microstrip patch antenna has been reported as small as this, that covers six application bands, as per the knowledge of the authors. HFSS ver.13 [24] software has been used to construct and analyze the proposed antenna before prototype fabrication.

II. ANTENNA DESIGN

The proposed ETMPA as shown in fig.1 with 15mm sides of the equilateral triangular radiating patch has been fabricated using square shape FR4 material with 22mm each side as substrate. The FR4 material has thickness of 3.2mm, relative permittivity (ε_r)=4.4 and tanδ=0.02.The ground plane has been made with same copper material as the radiating patch with same size as the size of the substrate. The equilateral triangular radiating patch of the antenna has been loaded with an I–slot of optimum length and breadth. The length, breadth and position of I-slot have been mentioned in fig.1 given below. To excite the proposed antenna a coaxial probe having 50Ω impedance has been used.

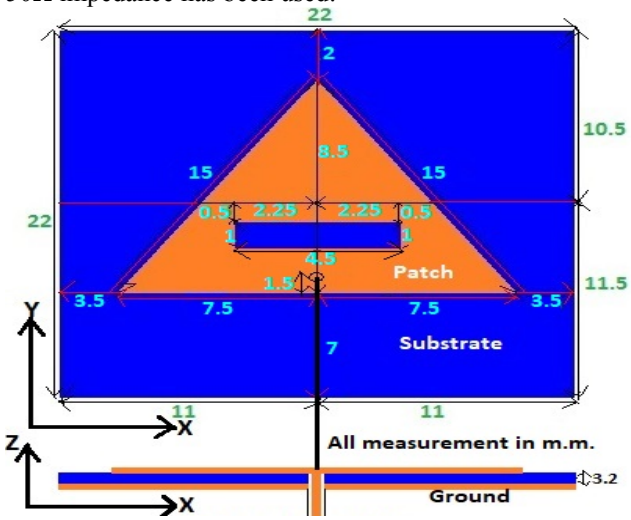


Fig.1 Design of proposed antenna

III. ANTENNA DESIGN PROCEDURE

The frequency response strongly depends on size & shape of the radiating patch, substrate material & height and ground size of the antenna [25]. In the process of designing the proposed antenna, first FR4 material having relative permittivity (ε_r) 4.4 has been selected as substrate because it is easily available and cheaper in cost. After that, theoretical calculation has been done using mathematical equation (1) & (2) [26] with equilateral triangular radiating patch having 15mm sides to get the primary resonant frequency (dominant mode TM_{10}) of the antenna, which has been found to be 5.82GHz after calculation.

$$a_{eff} = a + h\left[m^5 \left(0.0277 - 0.0083m\right) + \frac{0.07n\sqrt{\varepsilon_r}}{m^{0.55}} + \frac{8.1}{\varepsilon_r^2} - 1.05\sin\left(\sqrt{\frac{mn}{\varepsilon_r}}\right) + 1.2\frac{h}{a}\left(2.66\frac{n\sqrt{m}}{\varepsilon_r} - \frac{m^{3.765}}{\varepsilon_r^{2.1}}\right)\right] \quad (1)$$

$$f_{mn} = \frac{2c}{3a_{eff}\sqrt{\varepsilon_r}}\left[m^2 + mn + n^2\right]^{1/2} \quad (2)$$

Where,
a=Length of sides
h=Height of substrate
a_{eff}=Effective Length of sides
m,n=numbers to represent modes of resonant frequency
c=$3x10^{10}$cm/sec

After having calculated resonant frequency 5.82GHz, HFSS ver.13 software has been used to design and analyse the antenna, in order to find the probe position for optimal bandwidth and impedance matching. The optimum probe position has been located by shifting the probe in 1 mm steps over the median of the equilateral triangular patch from base to top vertex on the y-axis and analyzing result of each probe position. After analyzing results of all the probe position, it has been found that the probe position 1mm from base over median provided better results as compared to other probe positions. Further to fine tune the probe position, the probe has been shifted 0.5mm above and analysed the results, which is much better in comparison to all other probe positions. In this way the probe position has been optimised. A parametric analysis has been done with different probe position and graph of S_{11} with different probe position has been shown in fig.2.

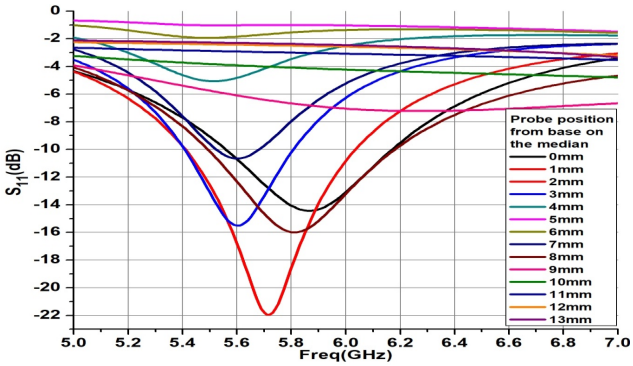
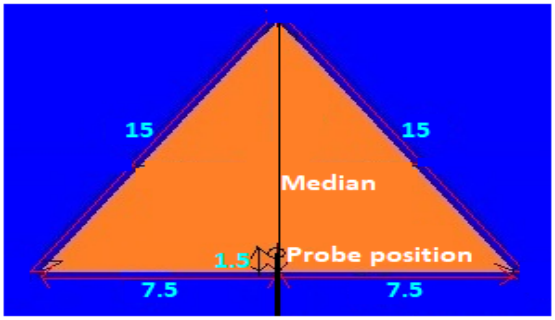
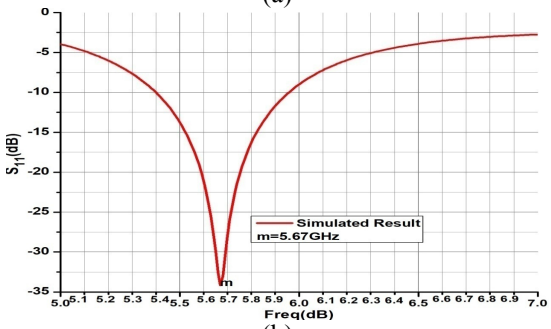


Fig.2. Graph showing $S_{11}$for different probe position on the median from base till the top corner of the triangular radiating patch.

At the probe position 1.5mm from base over median, primary resonant frequency is 5.67GHz, resonant frequency band(5.40-5.96GHz) for ($S_{11} \leq -10$ dB) with minimum S_{11} of -35.6dB as shown in fig. 3 below.



(a)



(b)

Fig.3.(a) Design of proposed antenna with optimum probe (1.5mm on the median from base) position. (b) S_{11} of antenna at optimum probe position

The bandwidth (5.40-5.96GHz) generated by antenna with optimum probe position does not cover desired application band. Therefore, equilateral triangular radiating patch of the proposed antenna has been loaded with an I-slot of optimum size and location to lower the resonant frequency band of the conventional antenna, so that the resonant band can occupy the desired application bands.

A. Effect of Slot

Slotting of radiating patch causes increasing surface current path due to which resonant band of the antenna shifts towards lower end of the frequency range [27-30]. In order to shift unslotted resonant band with optimum probe position, the patch has been loaded with an I-slot of dimension (1x1) mm everywhere on the radiating patch and been analysed at each point to find the desired result. After analysing the result of each point, it has been found that(1x1)mm slot at position 4mm above the base on the median and 0.5mm on the both sides of the median giving a result that can be modified by changing the length of I-slot to find the desired result. At above mentioned (1x1) mm slot position, the resonant band is (5.39-5.95) GHz, which shows a shift of 0.01GHz as compared to the unslotted result because of increasing surface current path. After having optimum I-slot position of dimension (1x1)mm, the length of the I-slot increases further on both sides of the median by 0.5mm, keeping breadth constant in order to facilitate further shifting of resonant band towards lower frequency range and results show little shift of resonant band. Finally with 1mm width and 4.5mm length of I-slot, the frequency band is shifted from (5.40-5.96) GHz to (5.34-5.95) GHz and occupied desired frequency band. In this way the dimension of I-slot on radiating patch has been optimises as mentioned in table 1 and fig.4 given below.

Tab.1. Showing shift of resonant band with increase of I-slot length.

Slot length (mm)	Resonant Band(GHz)	Minimum S_{11}(dB)
1	(5.39-5.95)	-31.2
2	(5.38-5.95)	-36.2
3	(5.37-5.96)	-42
4	(5.35-5.95)	-28.14
4.5	(5.34-5.95)	-23.7

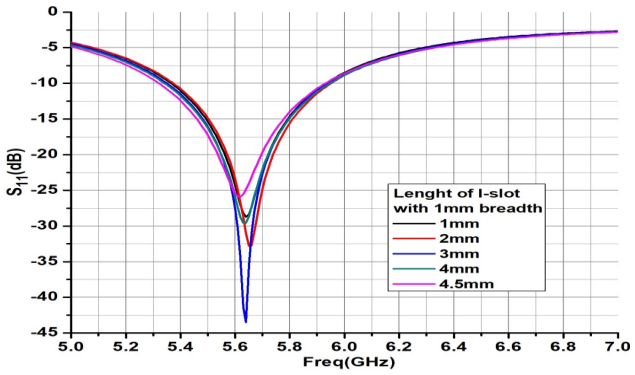
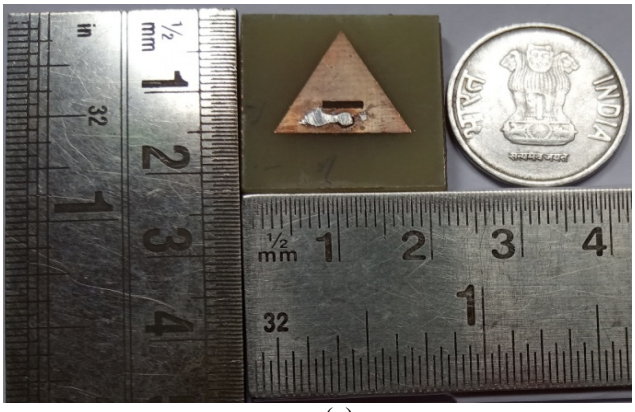
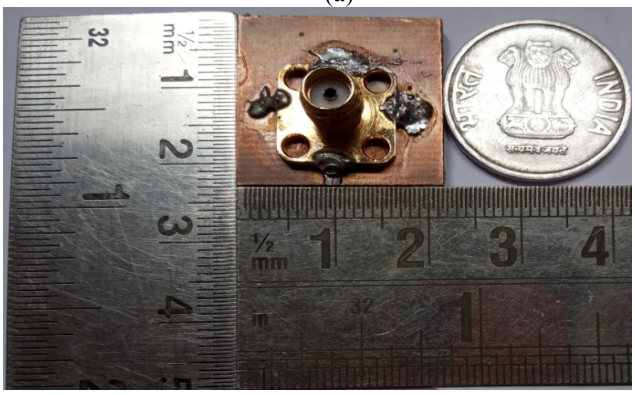


Fig.4$S_{11}$of proposed antenna showing variation of resonant bands with change in length of I-slot.



(a)



(b)

Fig.5. Photograph of fabricated prototype antenna: (a) Top view (b) Bottom view.

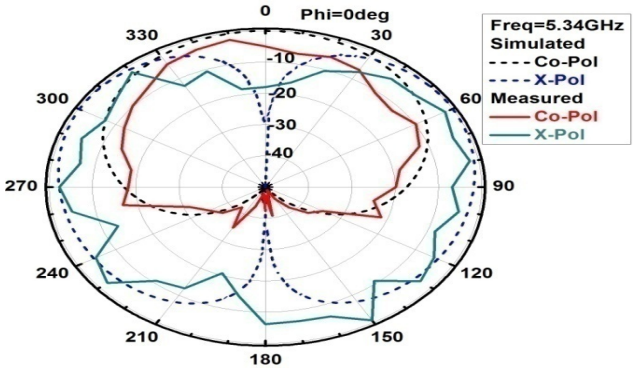
IV. RESULTS

The proposed antenna generated a wide band that covers six application bands (at probe position 1.5mm above base over median as shown in fig.1) having an I-slot of optimum size and position. Simulated and measured results of the proposed antenna have been provided in table 2 below.

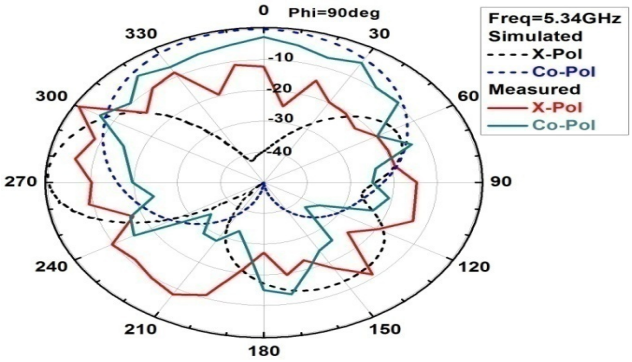
Tab.2. Shows simulated and measured results.

	Resonant Band with Minimum S_{11}(dB)	**Impedance Band width**	**Peak Gain at 5.64 GHz**	**Co-pol & X-pol difference**
Simulated	(5.34-5.95) GHz, -23.7 at 5.64 GHz	610MHz 10.8%	1.66 dB	39.6dB
Measured	(5.36-5.97) GHz, -17 at 5.66 GHz	610MHz 10.77%	1.32 dB	5.64GHz
Application Band	**1. Weather Radar application band (5.35-5.45) GHz [31].** **2. Radiolocation and Military application band (5.47-5.57) GHz [31].** **3. Maritime Radar application band (5.57-5.65) GHz [31].** **4. WLAN-Europe (5.470-5.725)GHz[32].** **5. Wi-Fi 5.8 (5.725– 5.850) GHz [33, 34].** **6. DSRC (5.850-5.925) GHz [35, 36].**			

Good matching has been found between measured and simulated results**.** But, due to fabrication tolerance[37], inaccurate soldering of the SMA (Sub-Miniature version A) connector, measuring environment [38] and inhomogeneous relative permittivity of substrate [39] which have not been considered at the time of simulation, little shift in the resonant frequency band and lower gain has been obtained in the measured result. At 5.34GHz, 5.64GHz and 5.95GHz of (5.34-5.95) GHz frequency range, the antenna's simulated & measured radiation pattern is broadside as shown in fig.6-8.

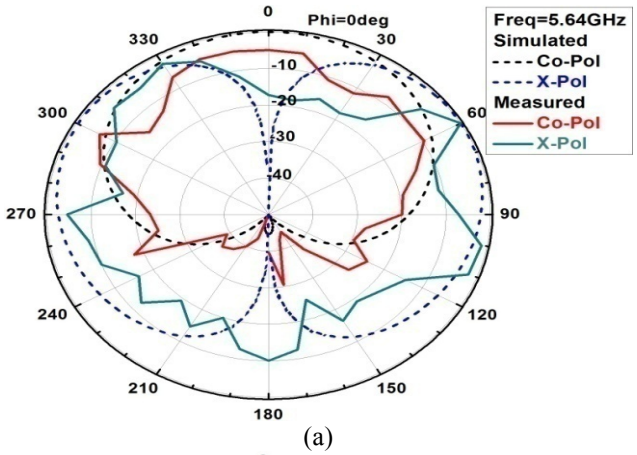


(a)

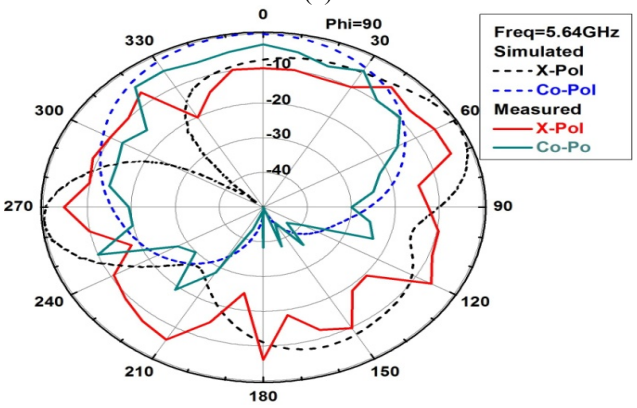


(b)

Fig.6.Radiation pattern (normalized)(a)E-Plane and (b)H-Plane at 5.34GHz.

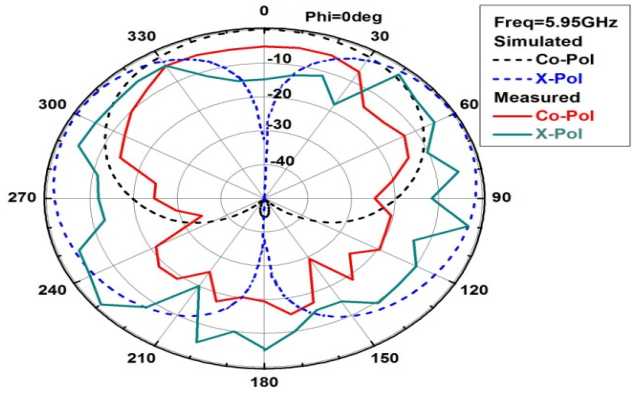


(a)

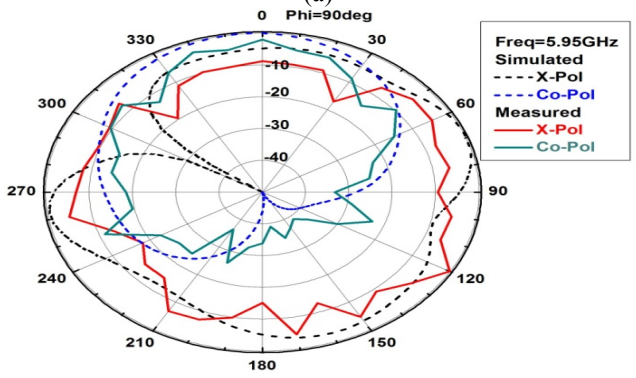


(b)

Fig.7.Radiation pattern (normalized) (a)E-Plane and (b)H-Plane at 5.64GHz.



(a)



(b)

Fig.8. Radiation pattern (normalized) (a)E-Plane and (b)H-Plane at 5.95GHz.

At above mentioned probe position with an I-slot S_{11}, Gain, Surface current distribution and Input Impedance of the proposed antenna are given in fig.9-12 below-

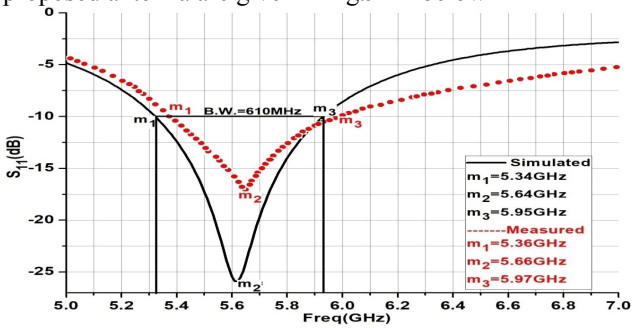


Fig.9. Simulated and measured S_{11} of the antenna.

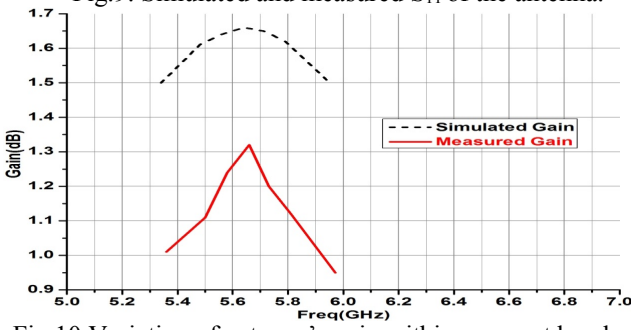


Fig.10.Variation of antenna's gain within resonant band with frequency.

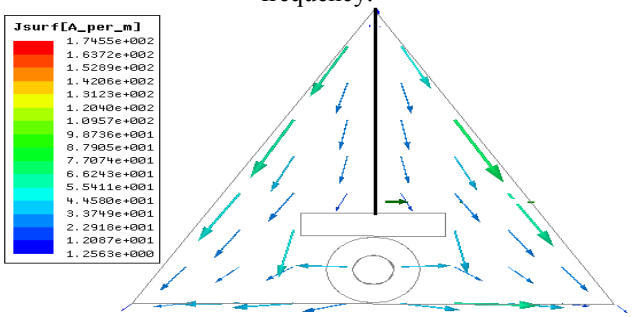


Fig.11. Simulated surface current distribution at 5.64GHz.

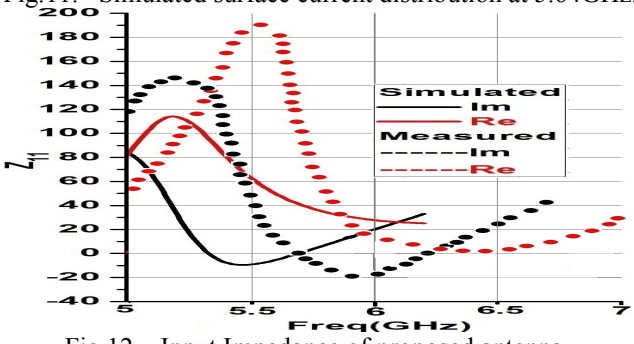


Fig.12. Input Impedance of proposed antenna.

A comparative study has been done between designed antenna performance and those available in literature, given below in Table 3.

Tab.3. Comparison of antenna performance with Existing design in literature.

Ref. no.	Side length of the patch(mm) & Area (mm^2)	Resonant Frequency (GHz)	Band width (%)	Gain
[16]	91 3585.8	Dual frequency 1.524, 2.342	4.07, 5.46	7dBi
[17]	48 997.7	Dual frequency 1.438, 1.962	1.67, 2.14	-
[18]	106 4865.3	Wide band 1.2	37.53	8dBi
[20]	28 339.5	Wide band 5.5	42	6dBi
[22]	91 3585.8	Broad band 1.677	18.3	-
[27]	28 339.5	Dual band 2.364, 4.995	4.2, 16.2	2.8 dBi, 5dBi
[29]	50 1082.5	Dual band 1.691, 3.268 Broad band 1.734, 1.784	1.66, 1.53 4.5	- -
[40]	50 1082.5	Broad band 1.76	9.3	6dB
[41]	136.8 8103.5	0.9	25.4	6dBi
[42]	31 416.1	Dual band 2.45, 5.8	21,27	2dBi, 3.5 dBi
My work	15 97.43	Wide band 5.64	10.8	1.32 dB

V. CONCLUSION

A very small equilateral triangular microstrip patch antenna having 15mm sides with an I-slot has been designed and analysed. The size and location of I-slot with optimum probe position has been optimized to obtain good results from the antenna. It generated wide frequency band, covering (5.34-5.95GHz) with good impedance matching and broadside radiation pattern for Weather Radar application band, Radiolocation and Military application band, Maritime Radar application band, WLAN-Europe, Wi-Fi 5.8 and DSRC band. The above designed antenna can be used in our advanced miniature communication equipments which uses above mentioned frequency bands and requires smaller space for antenna.

ACKNOWLEDGEMENT

Highly, thankful to Professor.(Dr.) Bhaskar Gupta of Jadavpur University, Kolkata for allowing me to use HFSS ver.13 software in his Lab. for my research work as well as Kaushik Patra Research Scholar of Jadavpur University for measurement of different parameters of fabricated antenna in the Lab.

REFERENCES

[1] Shaw, M., Chakravorty, D., Islam, S., Gangopadhyaya, M. "A simple star shaped microstrip patch antenna for penta band application." In *2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech),* Kolkata, 2018, p. 1-3.

[2] Shaw, M., Deb, B., Mandal, N. "Circular microstrip patch antenna with U-slots for multi band application." In *2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, Kolkata, 2018, p. 1-3.

[3] Xu, K.D., Xu, H., Liu, Y., Li, J., Liu, Q.H. "Microstrip Patch Antennas with multiple parasitic patches and shorting vias for bandwidth enhancement." In *IEEE Access*, 2018, vol. 6, p. 11624-11633.

[4] Khan, T., Roy, C. "Prediction of slot-position and slot-size of a microstrip antenna using support vector regression." *Int. J. RF Microw Comput.Aided Eng.*, 2018, vol.29, Iss.3.

[5] Jang, Y.-W. "Experimental study of a broadband U□slot triangular patch antenna." *Microw. Opt. Technol. Lett.*, 2002, vol.34, Iss.5, p.325-327.

[6] Balanis, C.A., Theory, A.'*Analysis and design*'(Wiley, New York, NY, USA, 2005.

[7] Sharma, V., Saxena, V.K., Saini, J.S., Bhatnagar, D., Sharma, K.B., Pal, D., Joshi, L.M. "Wideband dual□frequency right triangular microstrip antenna with parallel narrow slits." *Microw. Opt. Technol. Lett.*, 2010, vol.52, Iss.5, p.1082-1087.

[8] Toktas, A., Akdagli, A., Yagoub., M.C.E. "Calculating the dual-frequencies of equilateral triangular compact microstrip antennas with a shorting-Pin." *Microw.Opt. Technol. Lett*., 2013,vol.55, no.6, p.1227–1230.

[9] Shackelford, A.K., Leong, S.Y., Lee, K.F. "Small□size probe□fed notched patch antenna with a shorting post." *Microw. Opt. Technol. Lett.*, 2001, vol.31, Iss.5, p. 377 – 379.

[10] Chen, W., Lee, K.F. "Input Impedance of Coaxially Fed rectangular microstrip antenna on electrically thick substrate." *Microw. Opt. Technol. Lett*., 1993, vol.6, no.6, p.387-390.

[11] Lee, R.Q., Lee, K.F., Bobinchak,, J. "Characteristics of a two-layer electromagnetically coupled rectangular patch antenna." *Electron. Lett*., 1987, vol.23, no.20, p. 1070-1072.

[12] Hall, P.S. "Probe compensation in thick microstrip patches." *Electron. Lett.*, 1987, vol.23, p.606-607.

[13] Lee, K.F., Tong, K.F. "Microstrip patch antennas basic characteristics and some recent advances." *Proceedings of the IEEE*, 2012, vol.100, no.7, p.2169-2180.

[14] Lu, J.H. "Single-feed dual-frequency triangular microstrip antenna with a pair of bent slots." *Microw. Opt. Technol. Lett.,* 2001, vol.28, Iss.6, p.389-391.

[15] Lu, J.H. "Novel dual-frequency design of single-feed equilateral-triangular microstrip antenna." *Microw. Opt. Technol. Lett.,* 1999, vol.22, Iss.2, p.133-136.

[16] Wong, K.L., Pan, M.C., Hsu, W.H. "Single feed dual frequency triangular microstrip antenna with a V shaped slot." *Microw. Opt. Technol. Lett.*, 1999, vol.20, Iss.2, p.133-134.

[17] Wong, K.L., Fang, S.T., Lu, J.-H. "Dual-frequency equilateral-triangular microstrip antenna with a slit." *Microw. Opt. Technol. Lett.,*1998, vol.19, Iss.5, p.348-350.

[18] Deshmukha, A., Chaudhary, V., Shah, M., Pawar, S. "Wide band designs of equilateral triangular microstrip antennas with parasitic sectoral patches." In *8th International Conference on Advances in Computing and Communication (ICACC-2018)*, vol.143, p.87-93, 2018.

[19] Anguera, J., Puente, C., Borja, C., Romeu, J. "Miniature wideband stacked microstrip patch antenna based on the Sierpinski Fractal Geometry." In *IEEE Antennas and Propagation Society International Symposium*, vol.3, p.1700-1703, 2000.

[20] Mak, C.L., Luk, K.M., Lee, K.F. "Wideband triangular patch antenna." *IEE Proc-Microw. Antennas Prop*., vol. 146, no. 2, p.167-168, April 1999.

[21] Hassani, H.R., Syahkal, D.M. "Analysis of triangular patch antennas including radome effects." *IEE PROCEEDINGS-H*, vol. 139, no. 3, p.251-256, 1992.

[22] Wong, K.L., Hsu, W.H. "Broadband triangular microstrip antenna with U-shaped slot." *Electronics letters*, 1997, vol. 33 no. 25, p.2085-2087.

[23] Biswas, M., Dam, M. "Characteristics of equilateral triangular patch antenna on Suspended and Composite Substrates." *Electromagnetics*, 2013, vol.33, no.2, p.99-115.

[24] Ansoft High Frequency Structure Simulator (HFSS) ver.13.0, Pennsylvania: Ansoft Corporation.

[25] Mandal, K., Sarkar, p.p., et al. A compact low profile wideband U-shape antenna with slotted circular ground plane. *Int. J. Electron. Commun.(AEÜ),*2016, vol.70, p.336–340.

[26] Guney, K., Kurt, E. "Effective side length formula for resonant frequency of equilateral triangular microstrip antenna." *Int. J. of Electronics,* 2015, vol.103, no.2, p.261–268.

[27] Wu, J. "2.4/5GHz dual band triangular slot antenna with compact operation." *Microw. Opt. Technol. Lett*., 2005, vol.45, Iss.1, p.81-84.

[28] Dasgupta, S., Gupta, B., Saha, H. "Compact equilateral triangular patch antenna with slot loading." *Microw. Opt. Technol. Lett.,*2014, vol.56, Iss.2, p.268-274.

[29] Lu, J.-H., Tang, C.-L., Wong, K.-L. "Novel dual-frequency and broad-band designs of slot-loaded equilateral triangular microstrip antennas." *IEEE Trans. on Ant. and Prop.,* 2000, vol. 48, no. 7, p. 1048-1054.

[30] Mandal, B., Parui, S.K. "A miniaturized wearable button antenna for Wi-Fi and Wi-Max application using transparent acrylic sheet as substrate." *Microw. Opt. Technol. Lett.,* 2015, vol.57, Iss.1, p.45-49.

[31] Saini, S.S., Bhatia, H., Goel, A., Sidhu, E. "High return loss microstrip patch antenna design for radio applications." In *International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), Pune* , 2016, p. 1106-1109.

[32] Chakraborty, U., Kundu, A., Chowdhury, S.K., Bhattacharjee, A.K. "Compact dual-band microstrip antenna for IEEE 802.11a WLAN Application*." IEEE Ant.Wireless Prop. Lett.,* 2014, vol. 13, p. 407-410.

[33] Chaimool, S., Chung, K.L. "CPW-fed mirrored-L monopole antenna with distinct triple bands for WiFi and WiMAX applications." *Elect. Lett.,* 2009, vol.45, no.18, p. 928-929.

[34] Basar, M.R., Hossain, M.A., Hoque, M.R.U., Morshed, K.M. "Compact dual L-slit slotted antenna for spacecraft, WLAN, Wi-Fi, and Bluetooth application." In *7th International Conference on Electrical and Computer Engineering, Dhaka, 2012*, p. 706-708.

[35] Yem, V.V., Journet, B., Chi, P.V., Tu, V.T., Duc, N.V., Tien, P.V., Duc, N.T. "Novel high gain and broadband CPW-fed antennas with EBG for ITS applications." In *International Conference on Advanced Technologies for Communications (ATC), Ho Chi Minh City,* 2013, p. 451-456.

[36] Tiwari, N., Kumar, S. "Microstrip Patch Antenna for 5.9 GHz Dedicated Short Range Communication System." *Int. J. of Advance Electrical and Electronics Engineering (IJAEEE),* 2014, vol.3, Iss.4, p. 2278-8948.

[37] Islam, S.N., Kumar, M., Sen, G., Das, s. "Design of a compact triple band antenna with independent frequency tuning for MIMO applications." *Int.J. RF Microw Comput Aided Eng.,* 2019, vol.29, Iss.3.

[38] Gurjar, R., Upadhyay, D.K., Kanaujia, B.K.,Sharma, K. "A novel compact self-similar fractal UWB MIMO antenna." *Int J RF Microw Comput Aided Eng.,* 2019, vol.29, Iss.3.

[39] Zhang, C., Gong, J., Li, Y., Wang, Y. "Zeroth Order Mode Circular Microstrip Antenna with Patch-Like Radiation Pattern." *IEEE Ant.Wireless Prop. Lett.,* 2018,vol.99, p.1-1.

[40] Pan, M.C., Wong, K.L. "A broadband active equilateral triangular microstrip antenna." *Microw. Opt. Technol. Lett.,*1999, vol.22,no.6, p.387-389.

[41] Yeh, C.H., Hsu, Y.W., Sim, C.Y.D. "Equilateral Triangular Patch Antenna for UHF RFID Applications." *Int J RF and Microwave Comp Aid Eng,* 2014, vol.24, no.5, p.580-586.

[42] Liu, L., Zhu, S., Langley, R. " Dual-band triangular patch antenna with modified ground plane." *Electr. Lett*., 2007,vol. 43, no. 3.

A CPW-Fed Square Slot Circularly Polarized Antenna With Asymetric Stubs For Ultra Wideband Applications

Athul O Asok
Department of Electrical Engineering
IIT Palakkad
Kerala, India
athulasok14@gmail.com

Sukomal Dey
Department of Electrical Engineering
IIT Palakkad
Kerala, India
sdey28@iitpkd.ac.in

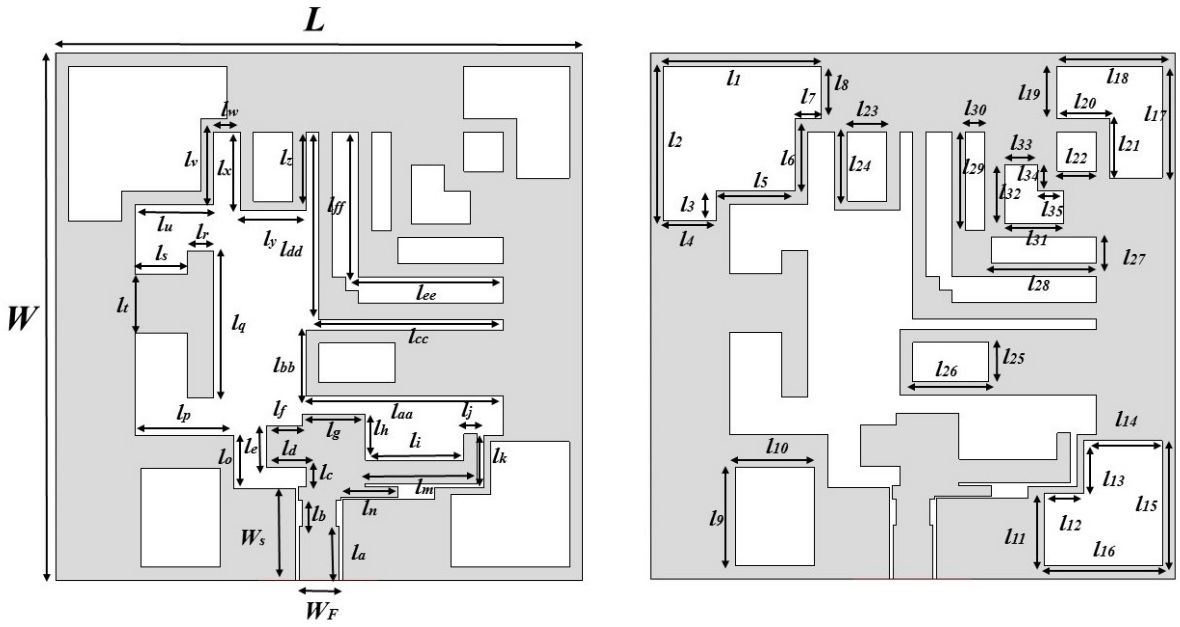
***Abstract*—This paper presents a new design for a Circularly Polarized Square Slot Antenna (CPSSA). The proposed antenna is single layered with L-shaped stubs loaded on both the opposite corners of the antenna, which is capable of generating orthogonal modes for the excitation of Circular Polarization. Two asymmetric strips on the top and the left portions of the proposed antenna and various other asymmetric strips are also loaded. The proposed structure is having the dimensions of 40 x 40 x 1.6 mm^3.The designed CPSSA operates in the range of 3.2-12.5 GHz with Axial Ratio (AR) bandwidth in the range of 4.5-8.5 GHz. In this paper the improvement of the performance parameters like AR and Return Loss (S_{11}) on the addition of asymmetric stubs is explained in detail.**

Keywords—Circularly Polarized Square Slot Antenna (CPSSA), Axial Ratio (AR), Return Loss (S_{11})

I. INTRODUCTION

Recently the rapid development of wireless communication has pushed the need for dual band, multiband and ultra-wideband (UWB) antennas. UWB antennas have large number of advantages compared to the dual and multiband antennas. UWB antennas have garnered a lot of attention since the FCC authorized the band from 3.1GHz to 10.6GHz as unlicensed [1].The devices that are currently used require huge data rate and large bandwidth to perform a variety of operations. Printed antennas with different kinds of feeding techniques (microstrip, co-planar and coaxial) have been used to achieve ultra-wideband characteristics. For seamless data transfer between the transmitter and receiver, without any mismatch, Circularly Polarized (CP) antennas are becoming more and more popular. Patch antennas are widely used in CP generation due to its compact size, low profile and symmetrical radiation. CP antennas are getting more recognition from mobile wireless communication. The advantages of CP include resistant to bad weather conditions and less sensitivity to the orientation of the wireless device. The advantages are more noticeable in satellite to land communication [2].Another reason CP antenna is gaining popularity is due to the fact that the CP operation eliminates the need to align the apertures of the two antennas which would otherwise be necessary to maximize the received power. A variety of design techniques have been proposed in the literature to obtain ultra-wideband CP antennas [3], [4].CP is a very attractive feature in ultra-wideband antennas. Bearing this in mind, various circularly polarized ultra-wideband antennas have been implemented in [5], [6].An asymmetric U-shaped slot was used in the ground plane to attain CP at 2.5GHz in [5].In [6], a rectangular stub was protruded out from the ground of the wideband antenna for getting an axial ratio bandwidth from 4.6-6.9GHz.In [7], a U-shaped slot in the ground plane and two feeds were used to achieve an axial ratio bandwidth (ARBW) from 1.83-6.35GHz and -10dB impedance bandwidth from 1.80-6.61GHz.To enhance the CP bandwidth many techniques are employed. One common technique is to increase the ARBW is to use thicker substrate with lower dielectric constant. But larger probe length could cause large inductance. To solve the issue, L-shaped probe [8] or L-shaped ground plane with air-layer acting as the antenna substrate [9] are suggested. In [10] a sequential rotation technique is explained which can increase the ARBW significantly. A novel microstip fed shorted square-ring slot antenna for circular polarization is elaborated in [11].The circularly polarized square slot antenna can provide broad impedance and axial ratio bandwidth [12].Right hand Circular Polarization (RHCP) and Left Hand Circular Polarization (LHCP) can be achieved with diverse techniques mentioned in [13], [14].Implanting a T-shaped grounded metallic strip which is perpendicular to the axial direction of the co-planar waveguide (CPW) feed is one of the method to increase the ARBW. This technique is elaborated in [15].Another popular method to improve the ARBW is by loading two L-shaped grounded strips on the opposite corners of the square slot antenna as given in [16]. Embedding a spiral slot in the ground plane [17] is often done for ARBW enhancement in square slot antennas. Inserting a lightning-shaped feed line with inverted ground strips [18] and utilizing the implanted arc shaped grounded metallic strip for circular and linear polarization [19] are also done for ARBW enhancement. A microstrip fed circular slot antenna for circular polarization is elaborated in [20].A new design of printed of a microstrip line fed circularly polarized ring slot antenna in which the CP is generated by introducing proper asymmetry in the ring slot structure and feeding the ring slot using a microstrip line at 45^0 from the introduced asymmetry is explained in [21].A novel broadband CP square-ring loaded slot antenna with flat gain, fed by an L-shaped microstrip line with tapered section for impedance matching is mentioned in [22].A circular slot antenna using L-shaped probe for broadband circular polarization is detailed in [23].In [24], a new design of a CP

978-1-7281-9615-2/20/$31.00 ©2020 IEEE



(a) (b)

Fig. 1.The geometry of the proposed CPSSA with asymmetric stubs. W=40mm, L=40mm, W_F=3mm, W_S=6.95mm, l_a=4mm, l_b=2mm, l_c=1.5mm, l_d=3mm, l_e=3.15mm, l_f=2.75mm, l_g=4.75mm, l_h=3.5mm, l_i=7.5mm, l_j=1mm, l_k=3.8mm, l_m=8.5mm, l_n=4.35mm, l_o=4mm, l_p=7.5mm, l_q=11.15mm, l_r=2mm, l_s=4mm, l_t=4.5mm, l_u=6mm, l_v=5.5mm, l_w=2mm, l_x=5.95mm, l_y=5mm, l_z=5.95mm, l_{aa}=15mm,l_{bb}=5mm, l_{cc}=14mm, l_{dd}=14.25mm, l_{ee}=11mm, l_{ff}=11mm, l_1=12mm, l_2=11.75mm, l_3=4mm, l_4=2.25mm, l_5=6mm, l_6=5.5mm, l_7=2mm, l_8=4mm, l_9=7.5mm, l_{10}=6mm, l_{11}=5.5mm, l_{12}=3mm, l_{13}=4mm, l_{14}=6mm, l_{15}=9.5mm, l_{16}=9mm, l_{17}=8.5mm, l_{18}=8mm, l_{19}=4mm, l_{20}=4mm, l_{21}=4.5mm, l_{22}=3mm, l_{23}=3mm, l_{24}=5.25mm, l_{25}=3mm, l_{26}=5.75mm, l_{27}=2mm, l_{28}=8mm, l_{29}=7.5mm, l_{30}=1.5mm, l_{31}=4.5mm, l_{32}=4.5mm, l_{33}=2.5mm, l_{34}=2mm, ,l_{35}=2mm .

square-ring slot antenna is proposed in which the CP is achieved by a series microstrip-line-feed configuration, coupling two orthogonal sides of the square-ring slot. Recently, metamaterial based CP antennas are becoming more popular. Metamaterials are those materials that show properties different from that of natural materials. A miniaturized CSRR loaded wide beam-width circularly polarized implantable antenna for subcutaneous real time glucose monitoring is explained in [25].Compact asymmetric-slit microstrip antenna for circular polarization is elaborated in [26]. Here, by cutting asymmetrical slits in the diagonal direction onto the square microstrip patches, the single coaxial-feed microstrip patch antennas are realized for CP radiation with compact size. Triple-band compact CP stacked microstrip antenna over reactive impedance meta-surface for GPS applications is detailed in [27]. A wideband CP antenna on wireless reactive substrates for telemetry applications is given in [28].Compact CP antennas combining meta-surfaces and strong space filling Meta resonators is given in [29].Finally, a single feed circularly polarized antenna loaded Complementary Split Ring Resonator (CSRR) is explained in [30] and diagonally asymmetric CSRRs loaded CP antenna with Frequency selective surface is elaborated in [31] –[34].

In this paper, a CPW-Fed Circularly Polarized Square Slot antenna with asymmetrical stubs is presented. Here, asymmetric stubs are used to generate circularly polarized Ultra-wideband antenna. The proposed structure has the dimensions of 40 x 40 x 1.6 mm^3 and is designed on an FR-4 substrate with dielectric constant 4.4 and loss tangent of 0.024. The structure achieves a good impedance bandwidth of 3.2-12.5GHz and a wide ARBW of 4.5-8.5GHz with a peak gain of 4dBi.

II. ANTENNA CONFIGURATION AND DESIGN

The geometry of the proposed CPSSA is shown in Fig.1.The antenna is designed on an FR-4 substrate with a dielectric constant of 4.4.The designed structure consists of various asymmetric stubs loaded on the square slot. Initially the T shaped stub is loaded on the CPW feed line on the structure. Then, an asymmetric T is loaded on the left portion of the slot. After this four L-shaped strips of various dimensions are loaded on the top right corner of the square slot. Once this accomplished, two slots are made on the two opposite corners, one being top left and the other being top right corners. The top left slot is initially a square of which two sides are extended to form an L shaped edge. The bottom right slot is also initially a square of which two adjacent sides are extended to finally obtain the desired structure. Similar modifications are also done on the bottom left and the top right corners as well. In the Bottom left corner, a square slot is etched out and in the corresponding top right corner two slots one square and then an L shaped slot is also created to achieve the desired performance. Two asymmetric strips, one U-shaped and the other one, a square strip with a rectangular slot are placed on the top as well as the right side of the square slot. The addition of these two strips brings down the operating frequency. Finally numerous modifications are done on the feed portions of the CPSSA. Thus, the antenna shows a wide operating frequency from 3.2-12.5GHz as shown in Fig. 2. Thus, the UWB region of operation is achieved. The addition of L-shaped strips on the corners of the square slot antenna and the creation of rectangular slots on the asymmetrical T shaped feed structure produces orthogonal modes which generates CP. Thus a good Axial Ratio Bandwidth (ARBW) in the range of 4.5-8.5GHz is obtained

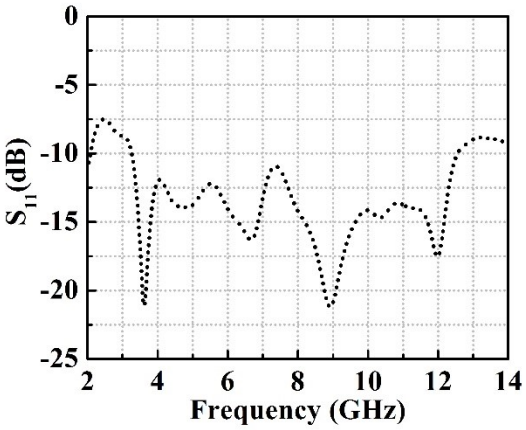


Fig. 2. The Return Loss Plot.

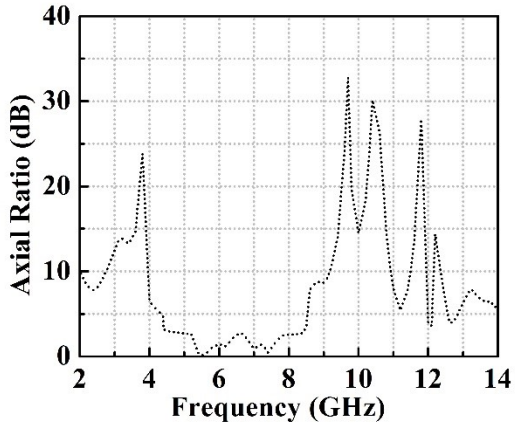


Fig. 3. The Axial Ratio Plot.

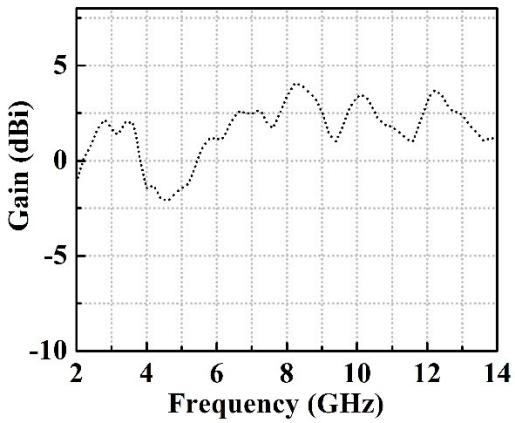


Fig.4.The Gain vs Frequency Plot.

as shown in Fig. 3. The gain of the CPSSA is plotted against frequency and the observations are noted from Fig. 4. It can be Inferred that a peak gain of 4dBi is obtained at 8.5GHz. The surface current distribution of the antenna at 7GHz is plotted in Fig. 5. The simulated results show that the current rotates in the clockwise (CW) direction. Thus the proposed antenna is Left Hand Circularly Polarized (LHCP).Simulated radiation pattern of the proposed antenna is shown in Fig. 6. The simulations are carried out in CST Microwave Studio 2016.

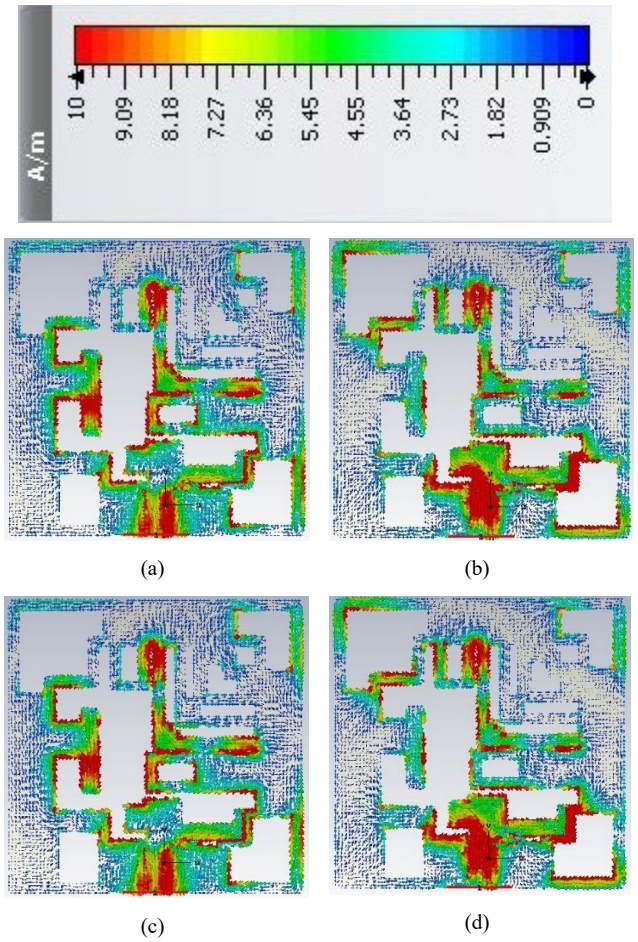


Fig. 5.The surface current distribution at 7GHz at (a)0^0 , (b)90^0 , (c)180^0 and (d)270^0

III. CONCLUSION

A new design for a Circularly Polarized Square Slot Antenna has been simulated in software. The proposed structure uses asymmetric stubs to achieve good impedance bandwidth as well as wide CP band. The gain, surface current and radiation pattern of the antenna has also been plotted and observed. The proposed design has a gain of 4dBi and is left hand circularly polarized. Parametric analysis shows that the CP characteristics can be varied by adjusting different design parameters of the antenna. Thus this antenna has the potential to be used in modern wireless communication systems.

ACKNOWLEDGEMENT

The authors would like to express their heartfelt gratitude to the Director for setting up the Central Instrumentation Facility(CIF) at the Indian Institute of Technology Palakkad. This work is supported by the Science and Engineering Research Board (SERB), Government of India under project no: ECR/2018/002258.

REFERENCES

[1] Federal Communications Commission, Government Printing Office, Washington: 'First report and order: Revision of part 15 of the commission's rules regarding ultra-wideband transmission systems', ET Docket 98-153, April 2002.

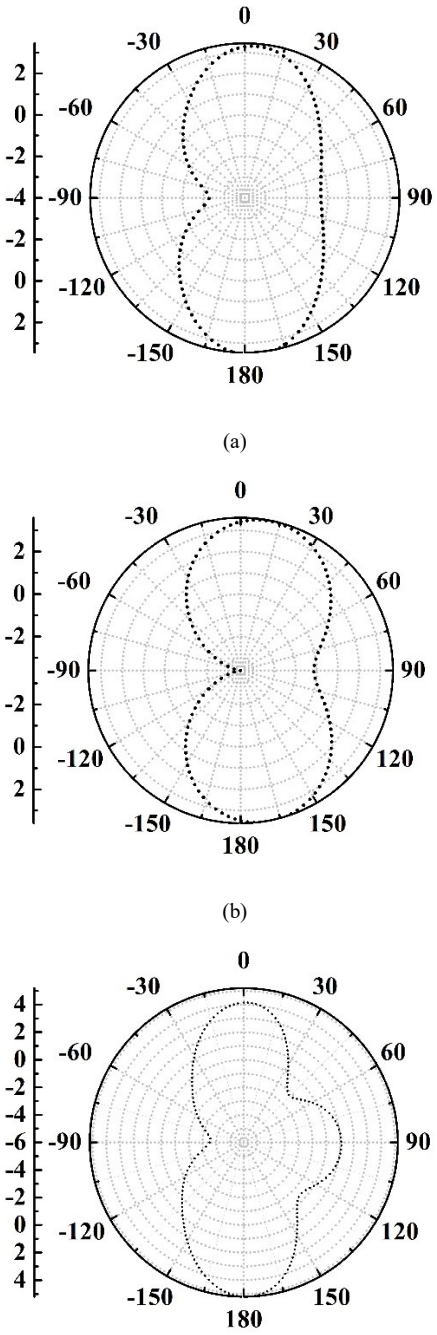


Fig. 5.Simulated Radiation Pattern at (a) 3GHz, (b) 5GHz and (c) 8GHz.

[2] P. Mousavi, M. Fakharzadeh, S. H. Jamali, K. Narimani, M. Hossu, H. Bolandhemmat, G. Rafi, and S. Safavi-Naeini, "A low-cost ultra low profile phased array system for mobile satellite reception using a zero-knowledge beamforming algorithm," *IEEE Trans. Antennas Propag*.,vol. 56, no. 12, pp. 3667–3679, Dec. 2008.

[3] A. K. Gautam, S. Yadav, and B. K. Kanaujia, "A CPW-fed compact UWB microstrip antenna," *IEEE Antennas Wireless Propag. Lett*., vol. 12, pp. 151–154, 2013

[4] A. K. Arya, R. S. Aziz, and S. O. Park, "Planar ultra-wideband printed wide-slot antenna using fork-like tuning stub," *Electron. Lett*., vol. 51, no. 7, pp. 550–551, Apr. 2015.

[5] S. Saxena, B. K. Kanaujia, S. Dwari, S. Kumar, and R. Tiwari, "A compact microstrip fed dual polarised multiband antenna for IEEE 802.11 a/b/g/n/ac/ax applications," *AEU-Int. J. Electron. Commun*., vol. 72, pp. 95–103, Feb. 2017

[6] M. S. Ellis, Z. Zhao, J. Wu, X. Ding, Z. Nie, and Q. H. Liu, "A novel simple and compact microstrip-fed circularly-polarized wide slot antenna with wide axial ratio bandwidth for C-band applications," *IEEE Trans. Antennas Propag*., vol. 64, no. 4, pp. 1552–1555, Apr. 2016.

[7] R. Xu, J. Y. Li, J. J. Yang, K. Wei, and Y. X. Qi, "A design of U-shaped slot antenna with broadband dual circularly polarized radiation," *IEEE Trans. Antennas Propag*., vol. 65, no. 6, pp. 3217–3220, Jun. 2017

[8] W. K. Lo, J.-L. Hu, C. H. Chan, and K. M. Luk, "Circularly polarized patch antenna with an L-shaped probe fed by a micostrip line," *Microwave Opt. Technol. Lett*., vol. 24, pp. 412–414, 2000.

[9] F. S. Chang, K. L. Wong, and T. Z. Chiou, "Low-cost broadband circularly polarized patch antenna," *IEEE Trans. Antennas Propag*., vol. 51, no. 10, pp. 3006–3009, Oct. 2003

[10] P. S. Hall, J. Huang, E. Rammos, and A. Roederer, "Gain of circular po-larized array composed of linearly polarized elements," *Electron. Lett.*,vol. 25, pp. 124–125, Jan. 1989.

[11] K. M. Chang, R. J. Lin, I. C. Deng, and Q. X. Ke, "A novel design of a microstrip-fed shorted square-ring slot antenna for circular polarization," *Microw. Opt. Technol. Lett*., vol. 49, no. 7, pp. 1684–1687, Jul. 2007.

[12] N. Felegari, J. Nourinia, C. Ghobadi, and J. Pourhmadazar, "Broadband CPW-Fed circularly polarized square slot antenna with three inverted Lshaped grounded strips," *IEEE Antennas Wireless Propag. Lett*., vol. 10, pp. 274–277, Apr. 2011.

[13] J.-Y. Sze, C.-I. G. Hsu, Z.-W. Chen, and C.-C. Chang, "Broadband CPW-fed circularly polarized square slot antenna with lightening-shaped feed line and inverted-L grounded strips," *IEEE Trans Antennas Propag*., vol. 58, no. 3, pp. 973–977, Mar. 2010.

[14] M.-J. Chiang, T.-F. Hung, and S.-S. Bor., "Dual-band circular slot antenna design for circularly and linearly polarized operations," *Microw. Opt. Technol. Lett*., vol. 52, no. 12, pp. 2717–2721, Dec. 2010.

[15] J.-Y. Sze, K.-L. Wong, and C.-C. Huang, "Coplanar waveguide-fed square slot antenna for broadband circularly polarized radiation," *IEEE Trans. Antennas Propag*., vol. 51, no. 8, pp. 2141–2144, Aug. 2003.

[16] J. Y. Sze and C. C. Chang, "Circularly polarized square slot antenna with a pair of inverted-L grounded strips*," IEEE Antennas Wireless Propag. Lett*., vol. 7, pp. 149–151, 2008.

[17] C. Chen and E. K. N. Yung, "Dual-band dual-sense circularly-polarized CPW-fed slot antenna with two spiral slots loaded," *IEEE Trans. Antennas Propag*., vol. 57, no. 6, pp. 1829–1833, Jun. 2009.

[18] J.-Y. Sze, C.-I. G. Hsu, Z.-W. Chen, and C.-C. Chang, "Broadband CPW-fed circularly polarized square slot antenna with lightening-shaped feed line and inverted-L grounded strips," *IEEE Trans Antennas Propag*., vol. 58, no. 3, pp. 973–977, Mar. 2010.

[19] M.-J. Chiang, T.-F. Hung, and S.-S. Bor., "Dual-band circular slot antenna design for circularly and linearly polarized operations," *Microw. Opt. Technol. Lett*., vol. 52, no. 12, pp. 2717–2721, Dec. 2010.

[20]] L. Y. Tseng and T. Y. Han, "Microstrip-fed circular slot antenna for circular polarization," *Microwave Opt. Technol. Lett*., vol. 50, pp. 1056–1058, Apr. 20

[21]] K. L. Wong, C. C. Huang, and W. S. Chen, "Printed ring slot antenna for circular polarization," *IEEE Trans. Antennas Propag*., vol. 50, pp. 75–77, Jan. 2002.

[22] H. Zhang, Y. C. Jiao, L. Lu, and C. Zhang, "Circularly polarized square-ring-loaded slot antenna with flat gains," *IEEE Antennas Wireless Propag. Lett*., vol. 15, pp. 1073–1076, Oct. 2017.

[23] R. Joseph, S. Nakao, and T. Fukusako, "Circular Slot Antennas Using L-shaped Probe for Broadband Circular Polarization", *Progress In Electromagnetics Research C (PIER C)*, vol. 18, pp.153-168, Jan. 2011.

[24]] J. S. Row, "The design of a squarer-ring slot antenna for circular polarization," *IEEE Trans. Antennas Propag*., vol. 53, pp. 1967–1972, Jun. 2005.

[25] X. Y. Liu, Z. T. Wu, Y. Fan, and E. M. Tentzeris, "A miniaturized CSRR loaded wide-beamwidth circularly polarized implantable antenna for subcutaneous real-time glucose monitoring," *IEEE Antennas Wireless Propag. Lett*., vol. 16, pp. 577–580, 2017.

[26] Dec. 2005. [6] Nasimuddin, X. Qing, and Z. N. Chen, "Compact asymmetric-slit microstrip antennas for circular polarization," *IEEE Trans. Antennas Propag*., vol. 59, no. 1, pp. 2

[27] K. Agarwal, Nasimuddin, and A. Alphones, "Triple-band compact circularly polarised stacked microstrip antenna over reactive impedance meta-surface for GPS applications," *IET Microw. Antennas Propag.*, vol. 8, no. 13, pp. 1057–1065, Oct. 2014.

[28] L. Bernard, G. Chertier, and R. Sauleau, "Wideband circularly polarized patch antennas on reactive impedance substrates," *IEEE Antennas Wireless Propag. Lett.*, vol. 10, pp. 1015–1018, 2011.

[29] H. Xu, G. Wang, J. Liang, M. Qi, and Z. Gao, "Compact circularly polarized combining meta-surfaces and strong space-filling meta-resonators," *IEEE Trans. Antennas Propag*., vol. 61, no. 7, pp. 3442–3450, Jul. 2013

[30] Dey, Soumik, Santanu Mondal, and Partha Pratim Sarkar. "Single feed circularly polarized antenna loaded with complementary split ring resonator (CSRR)." *Progr. Electromagn. Res.,* vol. 78, pp. 175-184, 2019.

[31] Dey, Soumik, Ankita Indu, Santanu Mondal, and Partha Pratim Sarkar. "Diagonally Asymmetric CSRRs Loaded Circularly Polarized Antenna with Frequency Selective Surface." *Progr. Electromagn. Res.,* vol. 92, pp. 43-54, 2020.

[32] Dey, Soumik, Santanu Mondal, and Partha Pratim Sarkar. "Reactive Impedance Surface (RIS) based asymmetric slit patch antenna loaded with complementary split ring resonator (CSRR) for circular polarization." *J. Electromagn. Waves Applications*, vol. 33, no. 8 pp. 1003-1013, 2019.

[33] K. Ding, C. Gao, T. Yu, and D. Qu, "Wideband CP slot antenna with backed FSS reflector," *IET Microw., Antennas Propag*., vol. 11, no. 7, pp. 1045–1050, Jun. 2017.

[34] A. Sharma, B. K. Kanaujia, S. Dwari, D. Gangwar, S. Kumar, H. C. Choi, and K. W. Kim, ''Wideband high-gain circularly-polarized low RCS dipole antenna with a frequency selective surface,'' *IEEE Access*, vol. 7, pp. 156592–156602, 2019

Design of a 'U' Slot Substrate-Integrated Waveguide Cavity-Backed Self-Diplexing Antenna

Rintu Kumar Gayen, Payel Midya, Sutanu Ghosh, Kaushik Patra, Malay Gangopadhyay
Department of Electronics & Communication Engineeinrg
Institute of Engineering and Management
Salt Lake, Kolkata - 700091, India
E-mail id: rintu.kumar.gayen@iemcal.com, payel.midya@iemcal.com, sutanu.ghosh@iemcal.com, kaushik.patra@iemcal.com, malay.ganguly@iemcal.com

Abstract — In this paper, the simple design of a "U" slot substrate-integrated waveguide (SIW)-based single substrate self-diplexing antenna is proposed. Two longitudinal slots (similar length) and one transverse slot are designated for "U" slot representation respectively, which radiate at two separate resonant frequencies, being imprinted on the top surface of the antenna. Two different microstrip feed lines are used to feed the SIW antenna. A high degree of isolation is achieved between the two ports.

Keywords—SIW antenna, self-diplexing, S-parameters, radiation pattern, slot antenna.

I. INTRODUCTION

With the rapid development of modern wireless communications systems, the need for low cost, low profile and high-performance planar dual-band antenna has been enhanced. Operating on satellite transceiver systems with various communication standards and various uplink/downlink frequencies, the compact wireless handheld device requires dual-band antenna with high isolation. Recently, the growth of self- diplexing antenna has involved the notice of researchers as it eliminates the need for high-order diplexers, thus plummeting RF front-end systems more compact, efficient and low-cost components [1]. Micro-strip technology-based self-diplexing antennas is an attempt to understand using the various techniques due to the characteristics described above [1-8].

In [1], a promising compact dual-frequency printed antenna is designed with orthogonally polarized modes and high isolation between the two feeding ports. The patch conductor shape is a circular sector and high dielectric constant material is used to provide good isolation between the up and downlink bands. In [2], a novel approach to design multi-frequency self-diplexed single patch antennas is proposed. This approach is based on a square microstrip patch antenna loaded with split ring resonators (SRRs) and in [3], the designed antenna is a self-diplexed patch antenna based on metamaterials for active RFID systems. In [4] a dual-fed, self-diplexing planar inverted RF antenna and an associated RF front-end are described. It is shown that co-design of the antenna and front-end can be used to double the operational bandwidth, without significant size or performance penalties. Indeed, the use of two feeds allows the antenna to be self-diplexing, which results in improved overall efficiency. A dual-feed dual-polarized microstrip antenna with low cross-polarization and high isolation is experimentally studied and two different feed mechanisms are designed to excite a dual orthogonal linearly-polarized mode from a single radiating patch in [5]. A mode-based design method for dual-band and self-diplexing antennas using double T-stubs loaded aperture is presented in [6]. Design of self-diplexing substrate integrated waveguide cavity-backed slot antenna is presented in [7], where a novel design technique to realize planar self-diplexing slot antenna using substrate integrated waveguide (SIW) technology is presented. The proposed antenna uses a bowtie-shaped slot backed by SIW cavity, which is excited by two separate feedlines to resonate at two different frequencies in X-band (8-12 GHz). Another substrate-integrated waveguide (SIW)-based self-diplexing antenna is presented [8]. Recently, a substrate integrated waveguide (SIW) cavity-backed self-triplexing slot antenna is proposed for triple-band communication [9]. A new design of self-triplexing slot antenna using substrate integrated waveguide (SIW) technique is demonstrated for multiband communication systems [10] and antenna-triplexer is realized using substrate integrated waveguide (SIW) technology

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

for radio altimeter/WLAN/ISM band applications [11].

In this communication, the simple design of a "U" slot substrate-integrated waveguide (SIW)-based single substrate self-diplexing antenna is proposed. Two longitudinal slots (similar length) and one transverse slot are designated for "U" slot represented respectively, which radiate at two separate resonant frequencies, are imprinted on the top surface of the antenna. Two different microstrip lines are used to feed the SIW antenna. A high degree of isolation is achieved between the two ports, which indicates self-diplexing feature.

II. Antenna Configuration

The geometry of the proposed antenna is shown in Figure 1. It is using a SIW cavity of dimension L × W acting as a resonator and of 1.57 mm height. The sidewalls of the space that contains the slots are formed by metallic vias, connecting upper and lower conducting planes. Two 50 Ω microstrip feed lines are used to feed the antenna from two opposite sides of the cavity. Two longitudinal slots and one transverse slot, designated as slot-I, slot-III and slot-II respectively, are etched on the top plane of the antenna. The lengths of slot-I, slot-III and slot-II are L1, L2 and L3, respectively, and each of the three slots are equally wide, Ws.

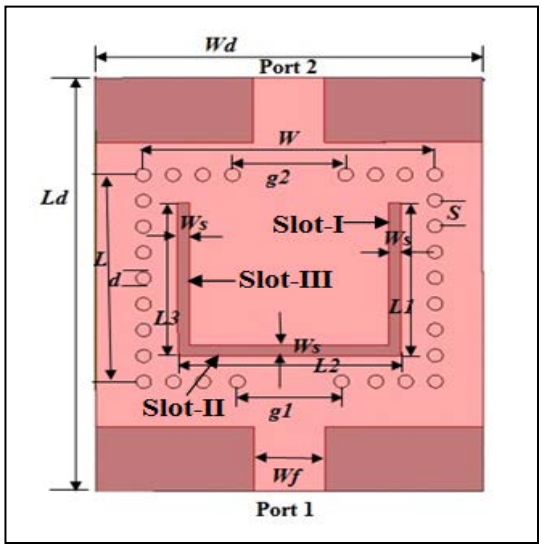


Figure 1: Geometry configuration of the "U" slot substrate-integrated waveguide cavity-backed self-diplexing antenna. [Ld = 32, Wd = 26, L1=L3 = 14.2, W = 20, L2 = 9.5, L2 = 11, Ws = 0.8,Wf = 4.8, d = 1, s = 2, g1 = 7 and g2 = 7.6 (all in mm)]. [13].

A. Principle of Operation

The slots have different wavelengths which are given two different resonant modes on simulation, with perturbed fields excited within the cavity. Slot-I and slot-3 are having longer slot length. When port 1 is excited, the antenna radiates at a comparatively lower frequency (F_L=19.5 GHz). When port 2 is excited, slot-II of the antenna radiates at a comparatively higher frequency (f_H =22 GHz). These happenings can be better explained with the help of surface current distribution on the top surface of the antenna at the two resonant frequencies. Figure 2(a) shows that at 19.5 GHz, on excitation of port 1, surface current density becomes dominant on the lower half of the antenna and is mostly concentrated around the boundary of slot-II. However, the surface current can almost be neglected in the vicinity of slot-I and Slot III. It is evident from the fact that, at this resonant frequency. The main actor behind the radiation is slot II and the role of slot-I and slot-III can be ignored. The radiation characteristics at second resonant frequency (f_H), located at 22 GHz, can be explained in a similar fashion, where the surface current is mostly concentrated along the boundaries of slot-I and slot-III [see Figure 2(b)]. In this case, port 2 is excited and radiation occurs mainly through slot-I and slot-III while slot-II has a negligible effect on it. The |S|-parameters of the proposed antenna are shown in Figure 3, where f_L and f_H are obtained at 19.5 GHz and 22 GHz, respectively, and the isolation is close to 10 dB in both the operating frequency bands.

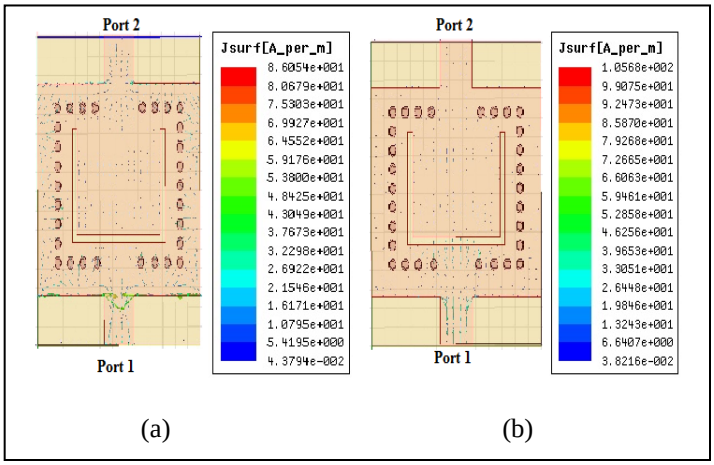


Figure 2. Surface current density distribution on the top plane of the proposed antenna at (a) 19.5 GHz (port 1 is excited) and (b) 22 GHz (port 2 is excited).

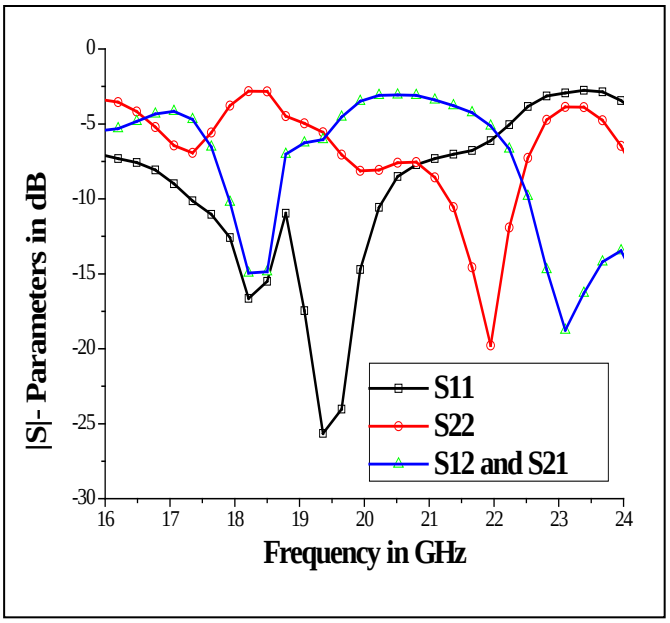


Figure 3. Variation of \S\-parameters with frequency for "U" slot substrate-integrated waveguide cavity-backed self-diplexing antenna.

The parametric study concludes that the isolation at f_H is not affected much and minimum isolation remains nearly 10 dB when the values of L1, L2 and L3 are changed. However, isolation at f_L varies highly between a wide range for differing values of L1, L2 and L3. Hence, by optimizing these parameters using full-wave electromagnetic simulator Ansoft HFSS 13.0, very good isolation, close to 10 dB, is acquired at both the operating bands. Hence, when one of the two ports is fed, very little energy is coupled to the other port. In this way, the proposed structure (antenna) is acting as an efficient diplexer as well as a dual-band antenna simultaneously.

III. Results and Discussion

To verify the proposed design, the proposed SIW-based self-diplexing antenna is simulated by Ansoft HFSS 13.0 on Rogers RT/duroid 5880 substrate (εr = 2.2, tanδ = 0.0009) of 1.57 mm thickness. The finally simulated dimensions of different parameters of the proposed antenna are shown in Figure 1.

The simulated |S|-parameters of the proposed antenna is plotted and shown in Figure 3, which display the simulated values. Two resonant frequencies of the simulated design are obtained at 19.5 and 22 GHz, respectively. The simulated minimum difference between reflection coefficients at two ports at first and second resonant frequencies are 25.6 and 20.2 dB, respectively.

The obtained radiation patterns on simulation are shown in Figure 4. The cavity-backed structure makes the radiation unidirectional and maximum radiation is toward the broadside direction at both of the frequencies.

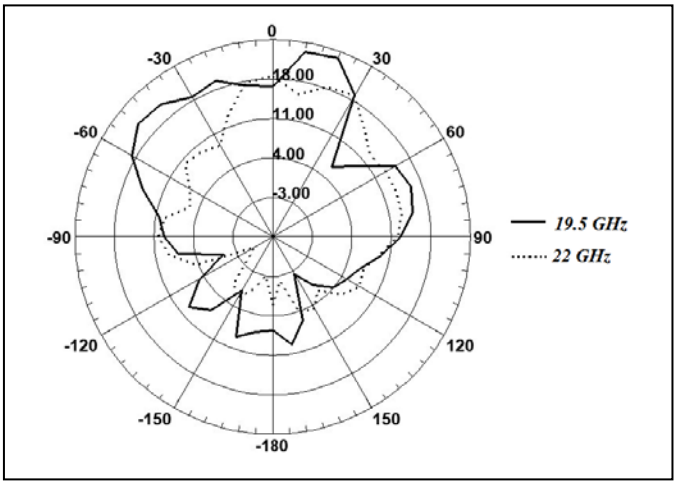


Figure 4. 2D radiation pattern of "U" slot substrate-integrated waveguide cavity-backed self-diplexing antenna (at f = 19.5 GHz and 22 GHz).

IV. Conclusion

A dual "U" slot-based SIW cavity-backed self-diplexing antenna is presented in this communication. Two longitudinal slots and one transverse slot are designated as slot-I, slot-III and slot II respectively. The proposed antenna offers ease in design and no trouble integrability with other RF modules due to its planar form and simple feeding method. The high isolation between two operating frequencies helps to get rid of multifaceted diplexers, most important to more solid RF front-end module. The lower and higher operating frequencies can be separately tuned from17.35 to 20.22 and 21.37 to 22.23 GHz, respectively. A satisfactory gain of 18 and 17dBi are achieved at 19.5 and 22 GHz i.e at resonant frequency, respectively, with unidirectional radiation pattern from the proposed antenna.

References

1. D. K. Park, R. Waterhouse, Y. Qian, and T. Itoh, "Self-diplexed integrated antenna transceiver for wireless applications," in Proc. IEEE

AntennasPropag. Soc. Int. Symp., 2001, vol. 3, pp. 444–447.

2. J.Montero-de-Paz et al., “Multi frequency self-diplexed single patch antennas loaded with split ring resonators,” Prog. Electromagn. Res., vol. 113,pp. 47–66, 2011.
3. F. J. Herraiz-Martinez et al., “Self-diplexed patch antennas based on metamaterials for active RFID systems,” IEEE Trans. Microw. Theory Techn.,vol. 57, no. 5, pp. 1330–1340, May 2009.
4. K. R. Boyle, M. Udink, A. de Graauw, and L. P. Ligthart, “A dual-fed selfdiplexing PIFA and RF front-end,” IEEE Trans. Antennas Propag., vol. 55,no. 2, pp. 373–382, Feb. 2007.
5. C. Y. D. Sim, C. C. Chang, and J. S. Row, “Dual-feed dual-polarized patch antenna with low cross-polarization and high isolation,” IEEE Trans.Antennas Propag., vol. 57, no. 10, pp. 3405–3409, Oct.2009.
6. Y. C. Lu and Y. C. Lin, “A mode-based design method for dual-band and self-diplexing antennas using double T-stubs loaded aperture,” IEEE Trans.Antennas Propag., vol. 60, no. 12, pp. 5596–5603, Dec. 2012.
7. S.Mukherjee and A. Biswas, “Design of self-diplexing substrate integrated waveguide cavity backed slot antenna,” IEEE Antennas Wireless Propag.Lett., vol. 15, pp. 1775–1778, 2016.
8. Sourav Nandi, and Akhilesh Mohan, “An SIW Cavity-Backed Self-Diplexing Antenna”, IEEE Antennas and Wireless Propagation Letters, Volume 16, PP. 2708 – 2711, 22 August 2017.
9. Kundan Kumar, and SantanuDwari, “A Self-Triplexing SIW Cavity-Backed Slot Antenna”, IEEE Antennas and Wireless Propagation Letters,Volume: 17, Issue: 5, May 2018.
10. Arvind Kumar andSingaravelu Raghavan, “A Self-Triplexing SIW Cavity-Backed Slot Antenna”, IEEE Antennas and Wireless Propagation Letters, Volume: 17, Issue: 5, May 2018.
11. D. Chaturvedi, A. Kumar, and S. Raghavan, “An Integrated SIW Cavity-Backed Slot Antenna-Triplexer”, IEEE Antennas and Wireless Propagation Letters, Volume: 17, Issue: 8, Aug. 2018.

Design Of Realistic Wearable Antenna Considering Superstrate And Human Body Perturbations

Sayantan Mukherjee
Department of Electronics and Communication Engineering
Institute of Engineering & Management
Saltlake, India
Email: dadaimukh@gmail.com

Sarbari Choudhury
Department of Electronics and Communication Engineering
Institute of Engineering & Management
Saltlake, India
Email: thisismesarbari@gmail.com

Saikat Choudhury
Department of Electronics and Communication Engineering
Institute of Engineering & Management
Saltlake, India
Email: thisismesaikat@gmail.com

Sayan Ghosh
Department of Electronics and Communication Engineering
Institute of Engineering & Management
Saltlake, India
Email: sayansonu14@gmail.com

Malay Gangopadhyay
Department of Electronics and Communication Engineering
Institute of Engineering & Management
Saltlake, India
Email: malay.ganguly@iemcal.com

***Abstract*—One of the latest trends in wireless technology is wearable antennas. This type of antenna needs to be flexible, lightweight and compact as they are meant to be a part of clothing and body. Properties of the textile antenna such as efficiency, input impedance, bandwidth etc. depend upon the type of substrate material used. These properties are mostly determined by the substrate dielectric constant. This paper describes the design of 2 types of rectangular microstrip patch antenna for on-body wireless communications in the 2.4GHz WLAN Band. In each design type, 3 different materials have been used as the dielectric substrate. The structure is simple, compact and easy to manufacture by using textile materials. The first design consists of three layers namely patch, substrate and ground plane. The second design consists of five layers namely superstrate, patch, substrate, ground plane and a tissue layer having properties similar to that of skin. The materials used as the substrate and superstrate for the 3 antennas are polyester, cordura and lycra. The patch and the ground plane in each of the 3 antennas have been made of copper. The simulation has been done using IE3D software. These antennas have a wide range of applications, especially in the field of healthcare.**

***Index Terms*—Wearable textile antenna, polyester, cordura, lycra, skin layer, IE3D software**

I. Introduction

The design of wearable antennas is such that it operates while it is worn on the body. Hence, these antennas need to be flexible, of low profile and also kept hidden for the comfort of the user. The best possible solution is to integrate the antenna with our clothing. Fabric or textile materials can be used to design such antennas as they are easily available and also used in our everyday life. In wearable context, antennas with microstrip patch design is most popular because of its advantages such as light weight, low volume, low profile planar configuration, low cost, etc. But it's limitation is that it has a narrow bandwidth. A wearable textile microstrip patch antenna, in general, consists of a radiating patch and a ground plane, both of which are made of conducting materials called electro-textiles. In between the patch and the ground plane, is a dielectric substrate which is also made of textile material and acts as an insulator. The ground plane of the antenna efficiently guards the body tissues and radiates vertically to the planar structure [2, 3] and therefore they are commonly used for wearable applications.[1-8]

The conductivity of the radiating patch is an essential parameter and it must be as high as possible[4]. The thickness and permittivity of the dielectric material [2] used as the substrate governs the efficiency and bandwidth of the antenna. The 2.4GHz ISM unlicensed band [6] is mostly used for the design of wearable antennas because of its global availability. The radiation pattern of a wearable antenna is equally important. Unlike the conventional rigid antennas, factors other than return loss, radiation pattern, gain and efficiency are also needed to be considered for the desired performance of a wearable antenna. Knowledge of electromagnetic properties such as permittivity and loss tangent of the textile material is also important. Clothing integrated with wearable textile antenna also provides extra facilities such as detecting, stimulating, and communication apart from protecting the body from heat, cold, etc.

978-1-7281-9615-2/20//$31.00 ©2020 IEEE

The design of the first type of antenna described in this paper is similar to the one mentioned above. The design of the second type of antenna is however a bit different. From top to bottom, the layers of this antenna are superstrate, patch, substrate, ground plane and a tissue layer having properties similar to that of skin. The functions of the patch, substrate and ground plane are same as those of the first type. The bottommost tissue layer will help us to have an idea of the effect of the antenna on human body. Similarly the topmost superstrate layer will help us understand the effectiveness of the antenna since in practical application, the patch will not be visible but will be covered by a layer of cloth or textile material. In each design type, a comparative study of 3 antennas, designed using 3 materials (having different dielectric values) as substrate and superstrate (in second type), has been done so as to identify the most suitable dielectric material to be used. The materials used are polyester, cordura and lycra. However, in all the cases, copper has been used as the conducting material to design the patch and the ground plane.

II. Design and Geometry of Wearable Textile Antenna (Type 1)

First we have designed a wearable textile antenna having 3 layers : patch, substrate and ground plane. The geometry of our antenna developed for body wearable application consists of 0.5mm thick insulating substrate on a thin conducting ground plane of copper and above the insulating substrate there is a rectangular copper patch which has been fed at correct position for matching the impedance with the surrounding coaxial probe's impedance. There are various conductive materials like copper, silver, electro textile and conductive sewing thread. Though it has unstretchable properties, copper has been used because of its low cost and excellent conductivity for microwave frequency application [9-14].

A. Antenna Design using IE3D software

The architecture of the proposed antenna consists of two conducting layers of copper (patch and ground plane) which have been separated via substrate. Finite ground plane size has been used (Length = 60mm and Breadth = 70mm). Three textile materials polyester, cordura and lycra with different values of dielectric constant have been selected. The design specifications of the 3 antennas are listed in Table 1. Zeland Inc's IE3D software has been used to simulate the wearable textile microstrip patch antennas.

B. Simulation Results

The simulation results in each case were noted for almost the same frequency (around 2.4GHz). The values for the different parameters as obtained from simulation are listed in Table II. The radiation patterns for each material for ϕ = 0 degree and 90 degree are also shown.

TABLE I
Design Parameters of Type 1 Antenna

Substrate	Polyester	Cordura	Lycra
Substrate Length (mm)	60	60	60
Substrate Breadth (mm)	70	70	70
Substrate Height (mm)	0.5	0.5	0.5
Substrate Epsr	1.9	1.6	1.5
Substrate tan d	0.004	0.0098	0.0093
Patch Length (mm)	46.25	49	51.5
Patch Width (mm)	52	55	55
Probe Position from centre (x,y)	9.5,0	12,0	12.5,0

TABLE II
Simulation Results of Type 1 Antenna

Substrate	Polyester	Cordura	Lycra
S-Parameters (dB)	-22.2983	-31.5898	-31.7393
Z-Parameters (Ohm)	53.6225	48.0324	49.4623
Total Field Directivity (dBi)	7.74497	7.98299	8.03454
Total Field Gain (dBi)	4.78363	3.23222	3.71917

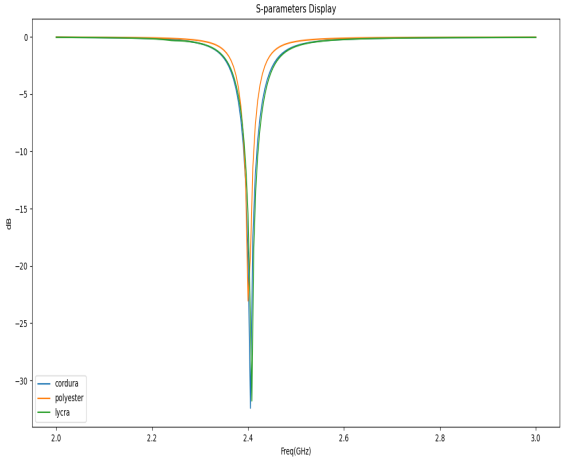


Fig. 1. S-parameter Plot for Type 1 Antenna

III. Design and Geometry of Wearable Textile Antenna Having Skin Layer and Superstrate Layer (Type 2)

The novelty of this design is that, in addition to the previous design of antenna, here we have included one layer of skin(Er=31.29,tan d=0.2835,height=1.5mm) below the ground plane and another layer of superstrate (Height=0.5mm and rest

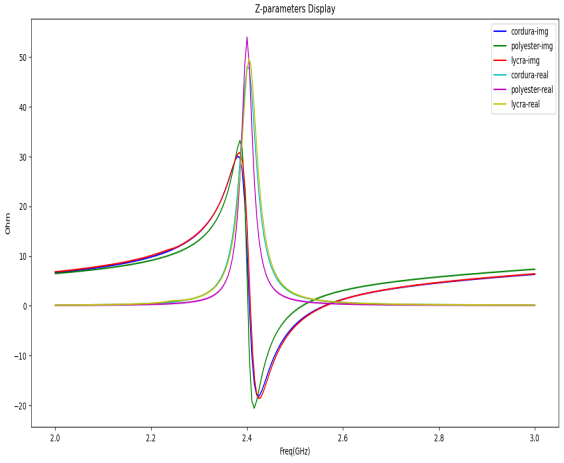


Fig. 2. Z-parameter Plot for Type 1 Antenna

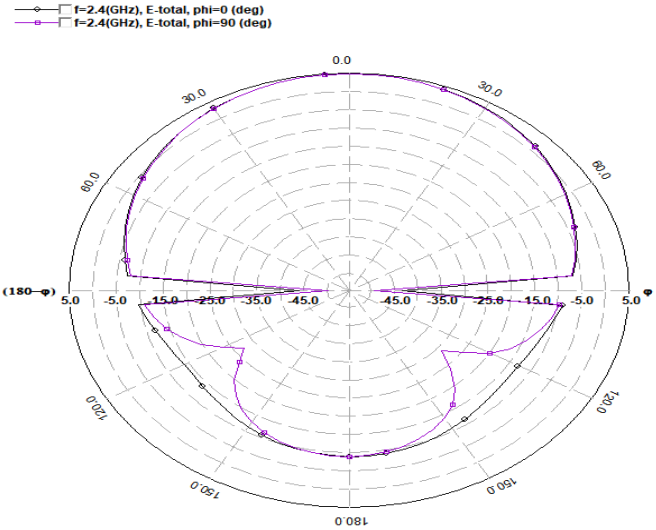


Fig. 4. Radiation Pattern for Type 1 Antenna using Polyester

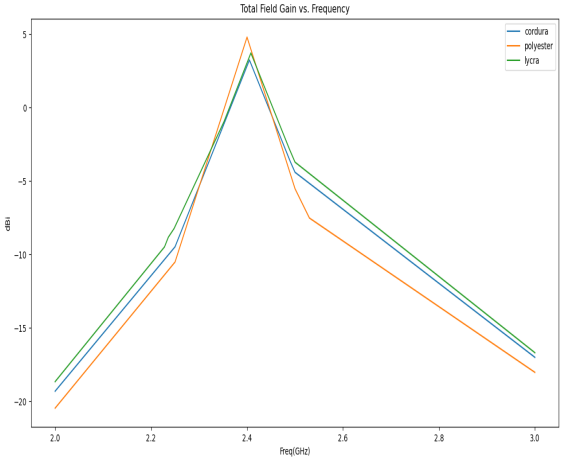


Fig. 3. Total Field Gain Plot for Type 1 Antenna

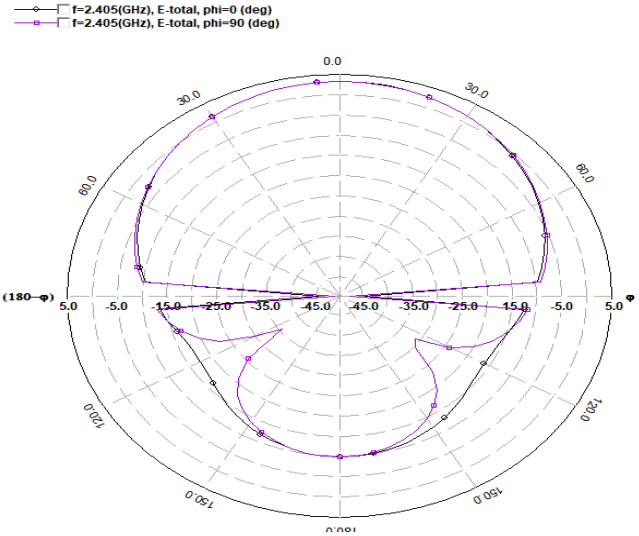


Fig. 5. Radiation Pattern for Type 1 Antenna using Cordura

same as substrate) above the Cu patch. Rest of the design has been kept similar to the previous antenna design.

A. Antenna Design using IE3D software

The architecture of the proposed antenna consists of two conducting layers of copper (patch and ground plane) which have been separated via substrate. On top of the patch, a superstrate layer made of the same material as substrate has been placed. Below the ground plane, a layer having the properties of skin has been included. Finite ground plane size has been used (Length = 60mm and Breadth = 70mm). Three textile materials polyester, cordura and lycra with different values of dielectric constant have been selected. The design specifications of the 3 antennas are listed in Table 3. Zeland Inc's IE3D software has been used to simulate the proposed wearable textile microstrip patch antennas. The top and side views of the two proposed antennas are shown in Fig.7 and Fig.8 respectively.

B. Simulation Results

The simulation results in each case were noted for almost the same frequency (around 2.4GHz). The values for the different parameters as obtained from simulation are listed in Table IV. The radiation patterns for each material for ϕ = 0 degree and 90 degree are also shown.

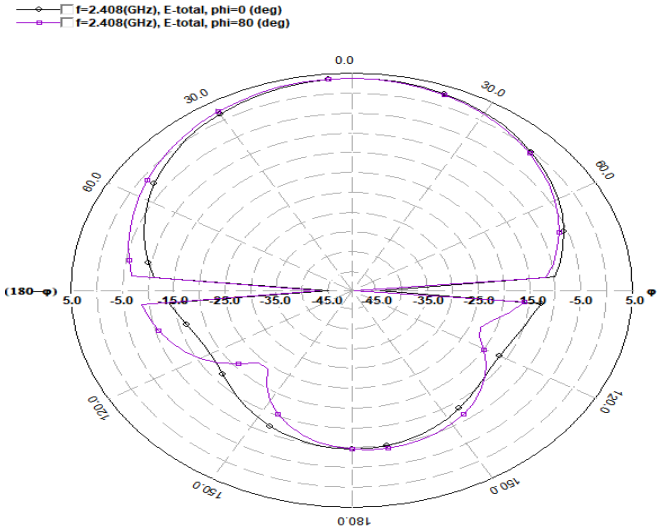


Fig. 6. Radiation Pattern for Type 1 Antenna using Lycra

TABLE III
DESIGN PARAMETERS OF TYPE 2 ANTENNA

Substrate/Superstrate	Polyester	Cordura	Lycra
Substrate/Superstrate Length (mm)	60	60	60
Substrate/Superstrate Breadth (mm)	70	70	70
Substrate/Superstrate Height (mm)	0.5	0.5	0.5
Substrate/Superstrate Epsr	1.9	1.6	1.5
Substrate/Superstrate tan d	0.004	0.0098	0.0093
Patch Length (mm)	44.25	44.25	49.8
Patch Width (mm)	55	55	55
Probe Position from centre (x,y)	10,0	13,0	13,0
Skin Height(mm)	1.5	1.5	1.5
Skin Epsr	31.29	31.29	31.29
Skin tan d	0.2835	0.2835	0.2835

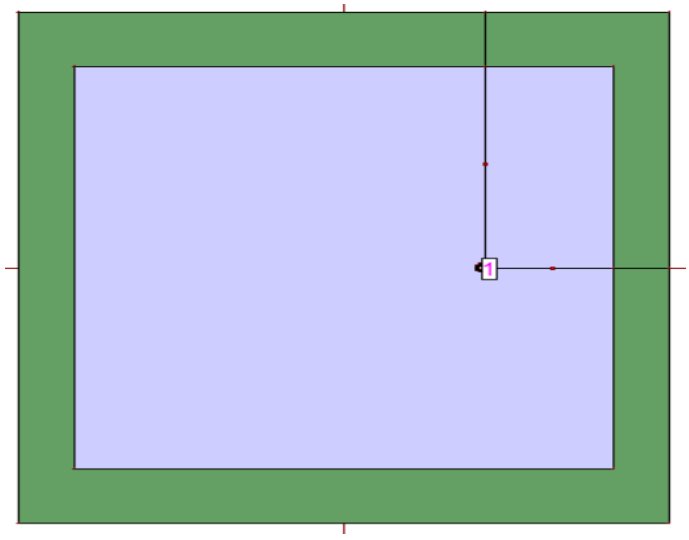


Fig. 7. Top View of the Proposed Antennas (Same for Both Types)

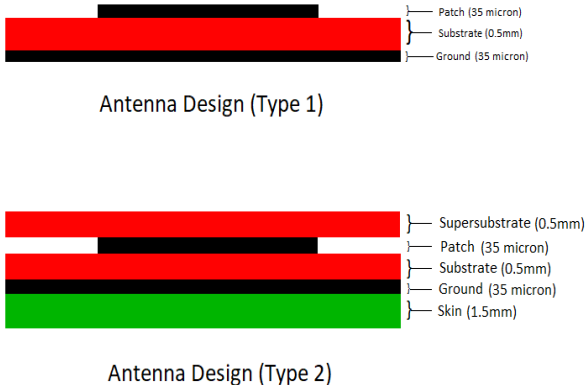


Fig. 8. Side Views of the Proposed Antennas

TABLE IV
SIMULATION RESULTS OF TYPE 2 ANTENNA

Substrate/Superstrate	Polyester	Cordura	Lycra
S-Parameters (dB)	-17.8921	-32.2683	-31.0568
Z-Parameters (Ohm)	50.0722	50.3399	49.4510
Total Field Directivity (dBi)	8.27747	8.27149	8.69334
Total Field Gain (dBi)	4.87161	3.00467	3.10556

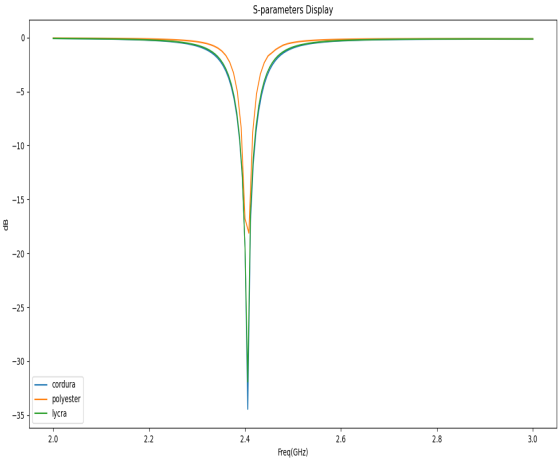


Fig. 9. S-parameter Plot for Type 2 Antenna

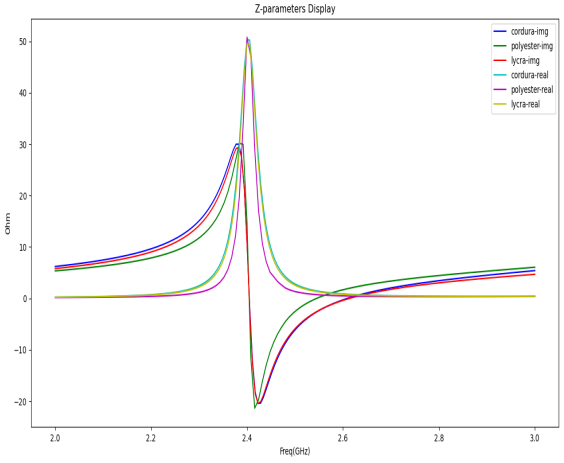


Fig. 10. Z-parameter Plot for Type 2 Antenna

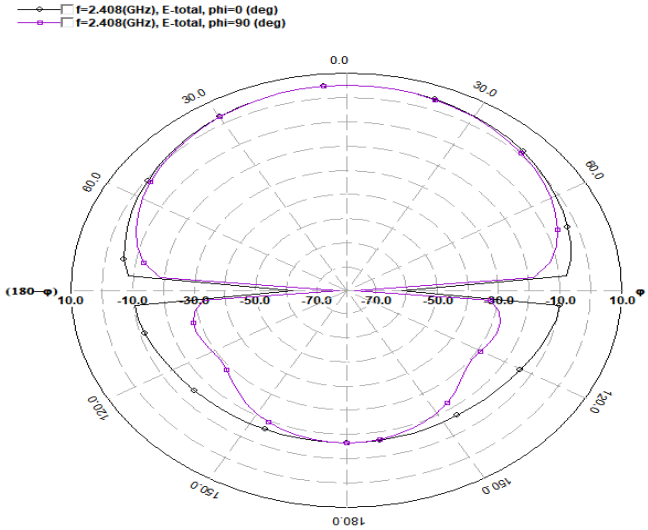


Fig. 12. Radiation Pattern for Type 2 Antenna using Polyester

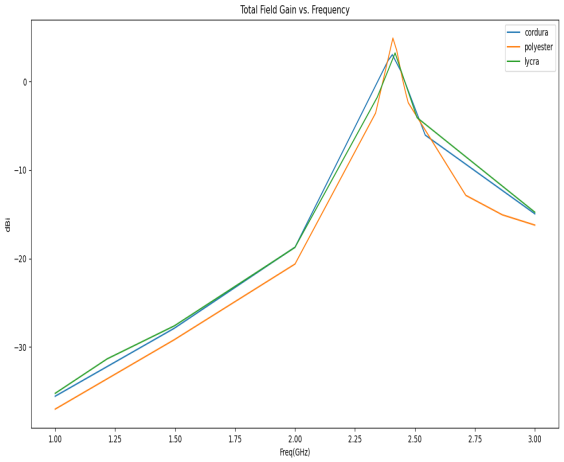


Fig. 11. Total Field Gain Plot for Type 2 Antenna

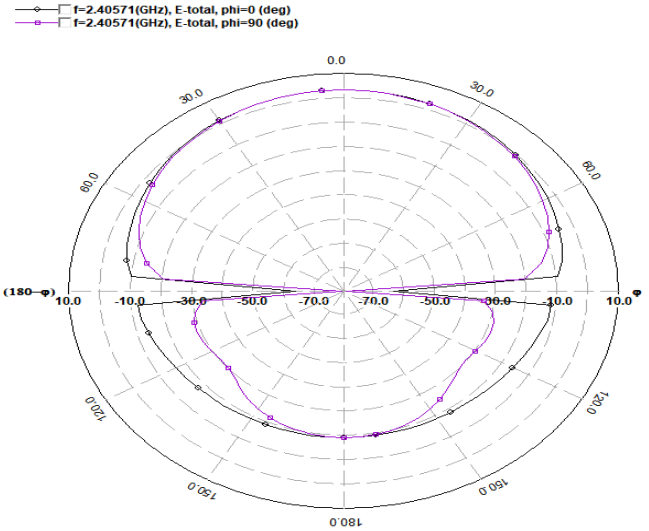


Fig. 13. Radiation Pattern for Type 2 Antenna using Cordura

IV. Analysis of Simulation Results

The impedance of the transmission line is approximately 50 Ohm (imaginary part = 0). Considering the antenna as the load, for maximum power transfer, the Z-parameter of the antenna should be same as that of the transmission line according to Maximum Power Transfer Theorem. S-parameter is given by 20logΓ, where Γ (Reflection Coefficient) is given by:

$$\Gamma = (Z_L - Z_0)/(Z_L + Z_0) \quad (1)$$

Z_0 is the characteristic impedance of transmission line and Z_L is the impedance of the antenna. As Z_0 becomes same as Z_L, Γ tends to 0 and hence the S-parameter tends to -∞ leading to better performance of the antenna. Thus, more negative the value of S-parameter, More effective is the antenna. An omnidirectional antenna radiates in all directions as a result of which some of the radiated power goes in vain. In case of wearable antenna, power radiated in the direction of human body has no utility. Gain of an antenna is a measure of the amount by which the designed antenna radiates in a specific direction with respect to an omnidirectional antenna. More the gain, more effective is the antenna.[15]

From Table II, it is observed that the S-parameters of antennas having cordura (-31.5898 dB) or lycra (-31.7393 dB) as substrate are better compared to that of the antenna designed using polyester substrate (-22.2983 dB). The Gain of the antenna designed using polyester (4.78363 dBi) however is the best among the three antennas (for cordura and lycra, the values are 3.23222 dBi and 3.71917 dBi respectively). These values mainly depend on the loss tangent values and

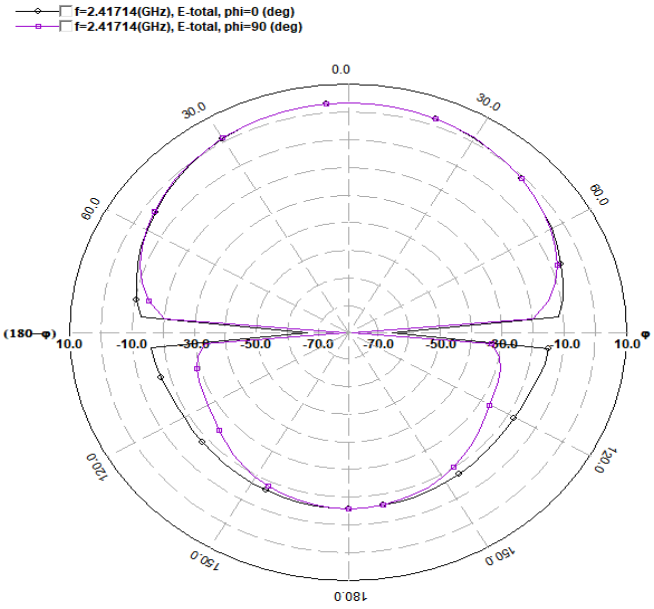


Fig. 14. Radiation Pattern for Type 2 Antenna using Lycra

permittivity values of the dielectric materials. From Table IV, it is observed that the S-parameters of antennas having cordura (-32.2683 dB) or lycra (-31.0568 dB) as substrate are better compared to that of the antenna designed using polyester substrate (-17.8921 dB). The Gain of the antenna designed using polyester (4.87161 dBi) however is the best among the three antennas (for cordura and lycra, the values are 3.00467 dBi and 3.10556 dBi respectively). These values mainly depend on the loss tangent values and permittivity values of the dielectric materials.

In each of Table II and Table IV, it can be seen that the Z-parameters in each of the cases is almost around 50 Ohm and the imaginary part is also zero (as can be seen from the graphs as well). Thus, according to Maximum Power Transfer Theorem, maximum power will be transmitted as the reflection coefficient (Γ) will tend to zero in each of these cases. Thus in both designs, the values obtained from simulation more or less satisfy the basic needs of antenna design.However, the current value of front to back ratio for the designed antenna can be attributed to the finite size of the ground plane. This can be subsequently mitigated using standard techniques available in literature, namely EBG (Electronic Band Gap) or Metamaterial Backing.

The second design is more accurate as it takes into consideration, the interaction between the human body and the antenna as well as the effect of the patch being covered by a textile material (as in practical application), on the working and effectiveness of the antenna. The results show that in both types of design, for polyester, the values of gain are highest and the values of S-parameters are also in the accepted range. Hence, polyester can be chosen as the dielectric material. However, working a little more on the other two materials can make them more suitable for use as well, since their values are also in the desired range.

V. Applications

A. In Healthcare monitoring

1) This patch antenna integrated into the textile can be used for wireless Body area aggregators and here Sensor nodes provide sensing various states of the wearer (like pulse rate, blood oxygen concentration ,ECG, Blood pressure) to the antenna wirelessly. Sensors, integrated into the clothing of the wearer, realize sensing functions and try to detect the state of the wearer (i.e. body temperature, pulse rate, position). Actuators enable actuating functions that provide signals (alarms) to inform and warn the wearer about certain conditions regarding his/her state or the state of the surrounding environment [16-19].
2) Our proposed patch antenna of 2.4GHz along with Toco sensors and Signal acquisition device can be implemented into the belly band or belt like wearable material to monitor the uterine contraction or infant respiration wirelessly [20].

VI. Conclusion

From the above analysis it can be concluded that with the help daily used fabric materials, this type of wearable textile antenna can be designed, which is low cost and can be easily integrated into fashion garments. The comparative study made using different dielectric materials can help us to select the most suitable material. From the results obtained from simulation, we can consider polyester to be the most suitable material, though the values for different parameters of the other two materials are also within permissible range. The first type of antenna designed gives us a basic understanding of the structure and requirements to design a wearable antenna. The second type of antenna, which is a modification of the first type helps us understand directly, the affects of the antenna on the body as it has a skin layer below the ground plane. In practical application, a wearable antenna is to be covered by a top layer of textile material to cover the radiating patch. This layer is also likely to affect the working of the antenna. The superstrate used on top of the patch in our second design considers that case as well. The second design can be further worked upon to improve the results by changing the design parameters. However, the effects of antenna bending on performance characteristics of wearable antennas requires further investigations. These wearable antennas must be flexible enough because the fabrics can take diverse shapes because of human body movements. They can prove to be useful in medical fields, in constant monitoring of human body, in diagnosis of diseases, and lot more. Proper work in the development of the second

type of antenna in particular can lead to a huge advancement in the field of wireless technology.

REFERENCES

[1] Ankita Priya, Ayush Kumar, Brajlata Chauhan, "A Review of Textile and Cloth Fabric Wearable Antennas" in International Journal of Computer Applications (0975 – 8887) Volume 116 – No. 17, April 2015.

[2] S. Sankaralingam and B. Gupta, "Determination of Dielectric Constant of Fabric Materials and Their Use as Substrates for Design and Development of Antennas for Wearable Applications," in IEEE Transactions on Instrumentation and Measurement, vol. 59, no. 12, pp. 3122-3130, Dec. 2010.

[3] L. Zhang, Z. Wang, D. Psychoudakis, J.L. Volakis, Flexible Textile Antennas for Body Worn Communication. In Proceedings of IEEE International Workshop on Antenna Technology, USA, pp. 205–208. March 2012.

[4] Pranita Manish Potey, Kushal Tuckley, " Design of Wearable Textile Antenna with Various Substrate and Investigation on Fabric Selection " in 2018 3rd International Conference on Microwave and Photonics,, 9-11 February, 2018.

[5] Salomen, P. and H. Hurme, "A novel fabric WLAN antenna for wearable applications," Proceedings of IEEE APS International Symposium, Vol. 2, 700–703, June 2003.

[6] S. Sankaralingam and B. Gupta, "DEVELOPMENT OF TEXTILE ANTENNAS FOR BODY WEARABLE APPLICATIONS AND INVESTIGATIONS ON THEIR PERFORMANCE UNDER BENT CONDITIONS", Progress In Electromagnetics Research B, Vol. 22, 53–71, 2010

[7] Locher, I., M. Klemm, T. Kirstein, and G. Troster, "Design and characterization of purely textile patch antennas," IEEE Transactions on Advanced Packaging, Vol. 29, No. 3, 777–778, November 2006.

[8] Tanaka, M. and J. H. Jang, "Wearable microstrip antenna, "Proceedings of IEEE APS International Symposium and URSI North American Radio Science Meeting, Columbus, OH, USA, June 2003.

[9] C. A. Balanis, Antenna Theory Analysis and Design, 3rd ed., Wiley-India, 2010, pp. 819 .

[10] Design and Characterization of Conformal Microstrip Antennas Integrated into 3D Orthogonal Woven Fabrics

[11] Titti kellomaki,effect of human body on single layer wearable antenna,Tampere University of technology ,publication 1025

[12] Rita Salvado, Caroline Loss, Ricardo Goncalves and Pedlo Pinho, Textile Materials for Design of Wearable Antennas: a survey,MPDI

[13] Morton W.E., Hearle W.S. Physical Properties of Textile Fibres. 4th ed. Woodhead Publishing; Cambridge, UK: 2008.

[14] R Garg, P Bhartia, IJ Bahl, A Ittipiboon , A Ittipiboon ,"Microstrip antenna design handbook", Artech house- 2001.

[15] Newman, E. H., P. Bohley, and C. H. Walter, "Two methods for the measurement of antenna efficiency," IEEE Trans. Antenna Propag., Vol. 23, No. 4, 457–461, July 1975.

[16] Kraus, J. D., R. J. Marhefka, and A. S. Khan, Antennas for All Applications, 3rd edition, 24–25, Tata McGraw-Hill Publishing Company Ltd., New Delhi, 2006.

[17] Maria Philip, Yao Xu, Projects:2016s2-230 New Materials for Wearable Antennas in Flexible Electronics.

[18] Laura Corchia, Giuseppina Monti, Egidio De Benedetto, and Luciano Tarricone, "Wearable Antennas for Remote Health Care Monitoring Systems" in International Journal of Antennas and Propagation 2017(6):1-11 ,December 2017.

[19] S. Majumder, T. Mondal, and M. Deen, "Wearable sensors for remote health monitoring," Sensors, vol. 17, no. 1, p. 130, 2017

[20] Dongmei Hao, Yang An, Xiangyun Qiao, Qian Qiu,Xiya Zhou and Jin Peng, "Development of Electrohysterogram Recording System for Monitoring Uterine Contraction" in Hindawi Journal of Healthcare Engineering, Volume 2019, Article ID 4230157, 9 pages.

Study on the CBOE Volatility Data Forecast Using Statistical and Computational Simulations

Richard Kyung
CRG-NJ
Cresskill, NJ, USA
richardkyung@choicerg.com

Minjun Kye
CRG-NJ
Cresskill, NJ, USA
minjunkye@choicerg.com

***Abstract*—**Economic indexes can be influenced by many different factors; therefore, it is difficult to use a single variable linear regression to determine the effectiveness of patterns. Modeling an economic pattern for a focused area and performing data analysis is especially difficult with a complex data pattern. To predict the effectiveness of such a trend, this paper focuses on a specific, objective main factor that determines the economic status in the field of stock markets. The CBOE Volatility Index, known by its ticker symbol VIX, is a popular measure of the stock market's expectation of volatility implied by S&P 500 index options. It is calculated and disseminated on a real-time basis by the Chicago Board Options Exchange (CBOE) and is commonly referred to as the fear index, or the fear gauge. In this paper, a statistical method is used to model the distribution of the maximum/minimum of a number of samples. Statistical measurements such as exceedance probability that an event exceeds mean value and return period are found based on historical data**.**

Keywords—Economic forecast, VIX, Gumbel distribution, Frechel distribution, Extreme value theory, Time series analysis

I. INTRODUCTION

Recent months have presented national economies with unprecedented challenges. With global lockdowns limiting normal consumption, plunging retail sales and supply shortages have become the new norm. Industries such as tourism, hospitality, and energy have largely ground to a halt, as oil prices approached 30-year lows. Stock markets across the world have seen frequent crashes since February 2020 and single week declines that exceed even the 2008 crisis, leaving the global economy in deep recession. Given the challenging and largely uncharted economic environment COVID-19 has created, it is crucial for businesses to fully understand the new risks involved, in order to successfully navigate these challenging times.

There are many factors that can determine economic index in the US, it is hard to use a single variable linear regression to determine the effectiveness of finding a pattern. In order to accurately predict the effectiveness of such a trend, an objective main factor that determines the status of economic status in the United States was chosen in this paper. The task of modeling economic patterns in a focused area and performing data analysis is not easy, especially when the data pattern is complex.

By using statistical and computational simulations, this paper investigates the impact of VIX on the economy and identifies risk factors associated with these outcomes. The VIX is one of the most widely used tools for economic index and stock selection. [2]

Extreme value theory is used to model the distribution of the maximum/minimum of a number of VIX samples of various distributions. [3] After finding exceedance probability that exceeds a critical value, a return period, which is an estimate of the likelihood of an event occurrence, was found. [4] A statistical measurement based on historic data denoting the average recurrence interval was found accordingly. Also, theoretical probability was calculated; and finally, estimated probability was checked to find if it approaches the theoretical probability as the number of trials gets larger.

II. BACKGROUND AND THEORY

A. Time Series Analysis and Forecasting

There have been numerous different kinds of data such as stock prices and interest rates observed and gathered in the past. The sequential nature of these data require us to account for the dynamic nature using special statistical skills and techniques. Time series analysis provides the appropriate methods necessary in order to analyze sequential data.

B. Smoothing

It may be problematic to picture the essential, underlying trend of the data if the time series has a lot of noise. To distinguish the signal and the noise from each other, various linear and nonlinear smoothers must be applied.

C. Curve fitting

In MATLAB®, best-fit lines are available using the least-sum-of-squares line from the data. Also known as linear least squares regression or least squares regression line (LSRL), this type of linear modeling minimizes the sum of the squares of the deviations between the model and the actual data. [6] The deviations are squared in order to reduce the influence of negative or positive signs when added.

In general terms, curve fitting involves either interpolation - in which the fitting model exactly matches the data, but often

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

in piecewise manners - or smoothing - in which the noise of the data is reduced and a function approximates the overall trend of the data. [7] [8] Curve fitting can be used to not only map out the data and render it computable in general terms, but also extrapolate other data points based on the trend provided by the model. [8]

III. DATA AND RESULTS

A. Plotting the VIX Frequency Curve using Gumbel Distribution

- Gumbel distribution is used to model the distribution of the maximum/minimum of a number of samples of various distributions. [3]
- Exceedance probability is the probability that the event will exceed some critical value (usually far from the mean).
- Return period is an estimate of the likelihood of an event to occur. [4] A statistical measurement based on historic data denoting the average recurrence interval.
- Theoretical probability is the fraction of times we expect the event to occur if we repeat the same experiment over and over (i.e. flipping a coin and getting heads or tails is each 0.50).
- Estimated probability approaches the theoretical probability as the number of trials gets larger. It is an approximation of theoretical probability.

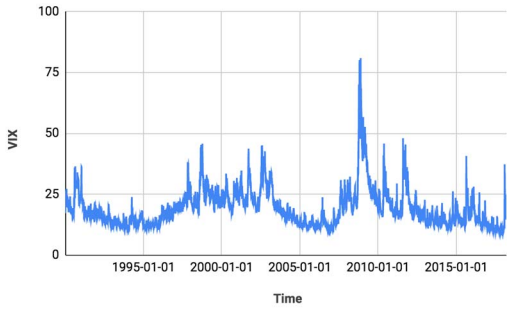


Fig.1. VIX data

1) Label the 1st column as 'time' and 2nd column as the 'VIX' and enter them in appropriate values in each cell.

2) Label 3rd column as 'Rank (i)' and rank the data in decreasing order (from N to 1).

3) Create a fourth column called q_i. Gringorten's plotting position formula will be used to calculate the estimated exceedance probabilities relevant to past observations. An example can be seen in Table C below.

$$q_i = \frac{i-a}{N+1-2a} \qquad (1)$$

q_i = exceedance probability associated with a specific observation
$N =$ number of annual maxima observations
$i =$ Rank of specific observation (i=1 is the largest and i=N is the smallest)
$a =$ constant for estimation (0.44)

4) Make another column and label it p_i. Then make it equal to $1 - q_i$. p_i refers to the non-exceedance probability.

B. Statistical Definition of Return Period

- If X is a random variable with a cumulative distribution function $F_x(x)$, the probability that X is less than equal (not exceeding) to a given event x_p is:
-

$$F_x(x) = P(X \leq x_p) = p \qquad (2)$$

- The probability that this event will be exceeded is now $1 - p$, and the percent exceedance would be $100(1 - p)\%$.
- For an event x_p, the return period corresponding to this exceedance probability is denoted by T.
-

$$T = \frac{1}{(1-p)} \qquad (3)$$

- For example, a 100-year return period is an event with a probability of exceedance $1 - p = 0.01$ or a non-exceedance probability $p = 0.99$. There is a 99% chance that this event will not be exceeded within a given year.

5) Create one more column and label it 'T_p estimated' and evaluate the values in p_i using the equation for the return period. An example can be seen in Table D below.

'T_p estimated' is the estimated distribution of 91 years of data. We assume that the data follows a specific distribution to estimate the parameters.

6) We will follow the 'Gumbel' or 'Extreme Value Type 1' distribution. The CDF (Cumulative distribution of function) of the Gumbel distribution is the following:

$$F_x(x) = exp\,[-exp\,(-\tfrac{x-u}{\alpha})\,] = p \qquad (4)$$

x is the observed VIXdata; u and α are the calculated parameters of the distribution. This distribution will allow us to calculate the theoretical estimate of p.

7) Create two columns labeled '$(x-u)/\alpha$' and 'p-theoretical. Using the following equations, calculate $\bar{x}$, s_x, u and α. Table E (below) shows such values that result from the existing data.

$$\underline{x} = \sum_{i=1}^{n} \left(\frac{x_i}{n}\right) \qquad (5)$$

$$s_x{}^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} [(x_i - \underline{x})^2] \qquad (6)$$

$$u = \underline{x} - 0.5772\alpha \qquad (7)$$

$$\alpha = \frac{\sqrt{6}s_x}{\pi} \qquad (8)$$

TABLE I. q_i, p_i, T_p ESTIMATED, AND $(x-u)/\alpha$OF VIX

time	VIX	rank	T_p estimated	$(x-u)/\alpha$
11/3/17	9.14	7053	0.003986707	-1.08875703
10/5/17	9.19	7052	0.003987272	-1.080613137
12/22/93	9.31	7051	0.003987838	-1.061067795
12/11/17	9.34	7050	0.003988404	-1.056181459
7/21/17	9.36	7049	0.003988969	-1.052923902
11/6/17	9.4	7048	0.003989535	-1.046408788
12/15/17	9.42	7047	0.003990101	-1.04315123
7/24/17	9.43	7046	0.003990668	-1.041522452
7/25/17	9.43	7045	0.003991234	-1.041522452
			continued...	
11/17/08	69.15	11	2.458041958	8.685543124
10/15/08	69.25	10	2.69348659	8.70183091
10/22/08	69.65	9	2.978813559	8.766982052
10/10/08	69.95	8	3.331753555	8.815845409
10/29/08	69.96	7	3.779569892	8.817474188
10/17/08	70.33	6	4.366459627	8.877738995
11/21/08	72.67	5	5.169117647	9.258873178
11/19/08	74.26	4	6.333333333	9.51784897
10/24/08	79.13	3	8.174418605	10.31106413
10/27/08	80.06	2	11.52459016	10.46254054
11/20/08	80.86	1	19.52777778	10.59284282

8) Use the VIX values (x) and populate the column $(x-u)/\alpha$ as shown in Table F below: Use the CDF equation from step 7 to calculate the value of p-theoretical.

9) Use the equation used to calculate 'T_p estimated' and use it to calculate 'T_p theoretical' using the p theoretical values, as seen in Table G below.

TABLE II. $(x-u)/\alpha$, P THEORETICAL, AND T_pTHEORETICAL OF VIX VALUES

(x-u)/a	p theoretical	$T_p$theoretical
-1.08875703	0.051273592	1.054044656
-1.080613137	0.052523972	1.055435674
-1.061067795	0.055606527	1.058880677
-1.056181459	0.056395254	1.059765759
-1.052923902	0.056925104	1.06036117
-1.046408788	0.057994501	1.061564928
-1.04315123	0.058534053	1.062173309
-1.041522452	0.058805044	1.062479132
-1.041522452	0.058805044	1.062479132
	continued...	
8.685543124	0.999831003	5917.253172
8.70183091	0.999833733	6014.413091
8.766982052	0.999844219	6419.271701
8.815845409	0.999851648	6740.7036
8.817474188	0.999851889	6751.690845
8.877738995	0.999860551	7171.059732
9.258873178	0.999904742	10497.7981
9.51784897	0.999926475	13600.82502
10.31106413	0.999966738	30063.911
10.46254054	0.999971413	34980.8074
10.59284282	0.999974905	39849.11074
-1.08875703	0.051273592	1.054044656
-1.080613137	0.052523972	1.055435674
-1.061067795	0.055606527	1.058880677
-1.056181459	0.056395254	1.059765759
-1.052923902	0.056925104	1.06036117
-1.046408788	0.057994501	1.061564928
-1.04315123	0.058534053	1.062173309

C. *Graphing the VIX Frequency Curve*

Go to 'insert' tab and select charts. Plot 'T_p estimated' vs Ratio Value. On the same graph, also plot 'T_p theoretical' vs Ratio Value. Label the chart title and the axes on the obtained graph. Figure 2 (below) shows how the graph may look like.

The Fisher-Tippett-Gnedenko Theorem, which provides a framework for analyzing a distribution that is continuous, has an inverse distribution. The following is the Fréchet law:

$$G(z) = exp\{-(\frac{z-b}{a})^{-a}\} \; z > b \qquad (9)$$

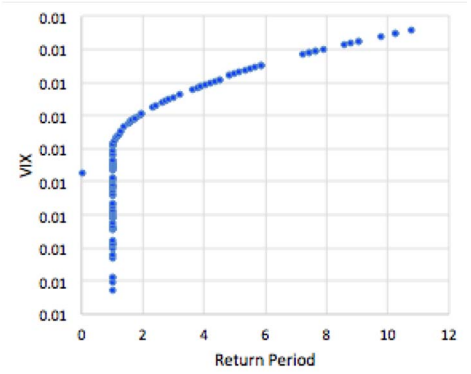


Fig.2. T_p estimated, T_p theoretical vs VIX Value

Right click on the curve and select 'change chart type' and click on 'scatter with smooth lines' for theoretical and 'scatter' for estimated. Right click again and select 'format chart area' to use the axis options command for the X-axis and select logarithmic scale. It will then contain return periods displayed from 1 to 100 in log scale. Figure 3 (below) shows what the graph may look like.

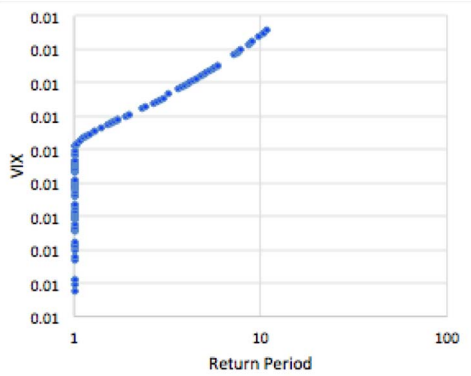


Fig.3. T_p estimated, T_p theoretical vs VIX Value, Logarithmic Scale

Fig. 3 shows the theoretical distribution, You can predict VIX corresponding to any return period from 1 to 100. The curve follows the distribution very well for low VIX, but starts to drift away from the theoretical at higher values.

We can see that lower VIX have much shorter return periods (meaning that they are more likely to repeatedly occur), while for higher VIX, the length of the return periods exponentially increase as the values increase (meaning that higher VIX are much less likely to repeatedly occur). This can be seen as a tendency for the VIX to remain in the lower ranges. The higher the VIX, the more over speculation and overestimation the stock market is, so it is natural for the VIX to come down after some time.

This trend is expected to repeat for some time in the future, since stock market bubbles are always bound to pop at some point, bringing the high VIX down to normalcy.

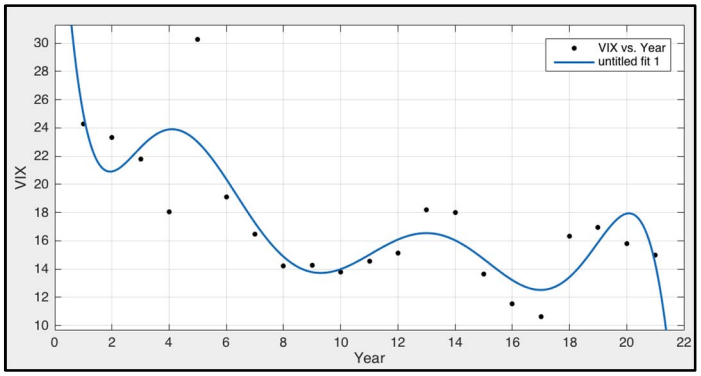


Fig.4. VIX with Polynomial Fit to 9 Degree

Figure 4 above shows the data with a polynomial model. When plotted in Matlab with the polynomial fit with 1 degree, the model was not a great fit to the actual data, since the R-squared value was nowhere close to 1 - in fact, it was closer to 0. Even if the degree was put to the maximum available in the software, the R-squared value was still much less than 1, let alone 0.5. This showed that a polynomial model would not suffice.

Linear model Poly8:

f(x) = p1*x^8 + p2*x^7 + p3*x^6 + p4*x^5 + p5*x^4 + p6*x^3 + p7*x^2 + p8*x + p9

Coefficients (with 95% confidence bounds):

p1 =2.39e-07 (-2.523e-06, 3.001e-06), p9 = 48.41 (-0.4245, 97.25)

Goodness of fit:

SSE: 126.4

R-square: 0.7003

Adjusted R-square: 0.5005

RMSE: 3.246

Fourier Series Model

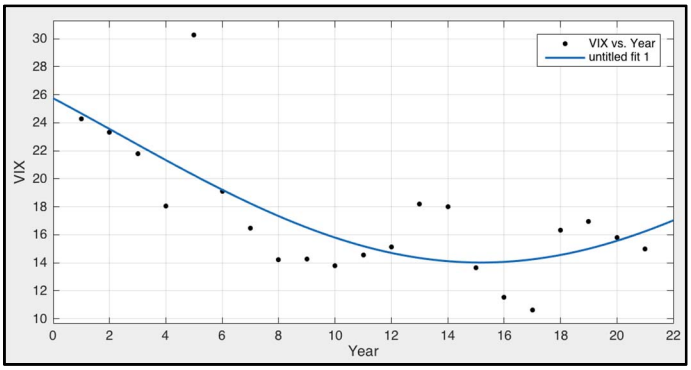


Fig.5. Polynomial - Fourier series (Matlab)

General model Fourier1:

f(x) = a0 + a1*cos(x*w) + b1*sin(x*w)

Coefficients (with 95% confidence bounds):

a0 = 23.12 (-22.21, 68.44)

a1 = 2.632 (-35.47, 40.73)

b1 = -8.711 (-65.99, 48.56)

w = 0.1228 (-0.2689, 0.5145)

Goodness of fit:

SSE: 192.4

R-square: 0.5438

Adjusted R-square: 0.4633

RMSE: 3.365

Holt's Linear Trend

Holt's Linear Trend uses the following equations:

$$u_1 = y_1 \quad (10)$$

$$v_1 = 0$$

$$u_i = \alpha y_i + (1-\alpha)(u_{i-1} + v_{i-1}) \quad (11)$$

$$v_i = \beta(u_i - u_{i-1}) + (1-\beta)v_{i-1} \quad (12)$$

$$\hat{y}_{i+1} = u_i + v_i \quad (13)$$

Where:

$0 < \alpha \leq 1$

$0 < \beta \leq 1$

Graphing Holt's Trend and Polynomials

Go to the 'insert' tab and select charts. Plot 't' vs VIX Value. Label the chart title and the axes on the obtained graph. Figure 2 (below) shows how the graph for the Holt's Trend may look like.

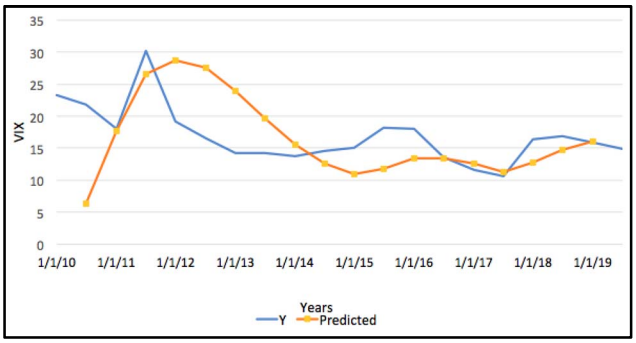


Fig.6. . y vs. time (Holt's method)

Overall, Holt's linearization works as a decent forecast of the VIX. There are discrepancies in multiple periods, however, a few notable periods such as 2012 and 2017 agree well. Data shows a tendency for the VIX to remain in the lower ranges.

IV. Discussion

Extreme value theory(EVT) assesses extreme events within probability distributions by quantifying tail behavior. By analyzing minimum or maximum values from samples, it seeks to determine probabilities for events more extreme than any previously observed events. Such analysis has applications in risk analysis involving hydrology, finance, insurance, etc.

This paper collected CBOE Volatility Index data and used the EVT(Extreme Value Theory) to map out the overall trend of the VIX in terms of its return period. Also, the data was plotted in Matlab, and multiple fitting models were tested out to see which one fit the data the best. The VIX was chosen due to its significance in the evaluation of stocks' values, and the Gumbel distribution due to its ability to incorporate rapidly fluctuating data into statistical analysis.

The Gumbel distribution was applied to distributions of maxima that have an exponential tail. The PDF for this Type is unbounded on both ends and is defined for all real values. The Frechet distribution was also used for the EVT in this paper. It is defined on an interval and converges to a certain value. The two types can be combined under the generalized extreme value distribution (GEV) to analyze the data better.

References

[1] Gumbel, E.J. (1935), "Les valeurs extrêmes des distributions statistiques" (PDF), Annales de l'Institut Henri Poincaré, 5 (2): 115–158

[2] Gumbel E.J. (1941). "The return period of flood flows". The Annals of Mathematical Statistics, 12, 163–190.

[3] Oosterbaan, R.J. (1994). "Chapter 6 Frequency and Regression Analysis" (PDF). In Ritzema, H.P. (ed.). Drainage Principles and Applications, Publication 16. Wageningen, The Netherlands: International Institute for Land Reclamation and Improvement (ILRI). pp. 175–224. ISBN 90-70754-33-9.

[4] ASCE, Task Committee on Hydrology Handbook of Management Group D of (1996). Hydrology Handbook | Books. doi:10.1061/9780784401385. ISBN 978-0-7844-0138-5.

[5] Peres, D. J.; Cancelliere, A. (2016-10-01). "Estimating return period of landslide triggering by Monte Carlo simulation". Journal of Hydrology. Flash floods, hydro-geomorphic response and risk management. 541: 256–271. doi:10.1016/j.jhydrol.2016.03.036.

[6] Anonymous (2014-11-07). "Flood Estimation Handbook". UK Centre for Ecology & Hydrology. Retrieved 2019-12-21.

[7] https://en.wikipedia.org/wiki/Price%E2%80%93earnings_ratio#In_business_culture

[8] https://en.wikipedia.org/wiki/Gumbel_distribution

[9] https://en.wikipedia.org/wiki/Return_period

[10] https://www.usna.edu/Users/oceano/pguth/md_help/html/time6h9j.htm

[11] https://www.itl.nist.gov/div898/handbook/pmd/section1/pmd141.htm

[12] http://web.iitd.ac.in/~pmvs/courses/mel705/curvefitting.pdf

[13] https://en.wikipedia.org/wiki/Curve_fitting

[14] A high-resolution rainfall re-analysis based on radar–raingauge merging in the Cévennes-Vivarais region, France Brice Boudevillain, Guy Delrieu, Annette Wijbrans, Audrey Confoland Pages 14-23

[15] Estimation of intensity-duration-frequency relationships according to the property of scale invariance and regionalization analysis in a Mediterranean coastal area

An Efficient Approach for Task Assignment in Spatial Crowdsourcing

Esam Aloufi
Computer Science and Engineering
Oakland University
Rochester, USA
aloufi@oakland.edu

Raed Alharthi
Computer Science and Engineering
Oakland University
Rochester, USA
rsalharthi@oakland.edu

Mohamed Zohdy
Computer Science and Engineering
Oakland University
Rochester, USA
zohdyma@oakland.edu

Dareen Alsulami
Department of Educational Technology
University of Jeddah
Jeddah, KSA
dmalsulami1@uj.edu.sa

Ibrahim Alrashdi
Computer and Information Science
Jouf University
Sakaka, KSA
irrashdi@ju.edu.sa

Richard Olawoyin
Industrial and Systems Engineering
Oakland University
Rochester, USA
olawoyin@oakland.edu

***Abstract*—Spatial crowdsourcing is a form of location-based crowdsourcing. With the spread use of mobile phones and smart devices that can share location, spatial crowdsourcing gained a lot of attention, especially in ride-hailing services. This paper evaluates the performance of a proposed spatial crowdsourcing task assignment approach to increase the task assignment rate while preserving the location privacy of the crowd workers. The overall experiments on real-world data sets show that the proposed approach results in the maximal total number of assigned tasks without significant disclosure of crowd workers' locations.**

***Index Terms*—Crowdsourcing, Spatial Crowdsourcing, Location privacy, Spatial task assignment.**

I. Introduction

Crowdsourcing is a computing paradigm which divides tasks permitted in crowdsourcing into smaller tasks for implementation, namely: Sequential Implementation, Parallel Implementation, Divide Conquer Implementation, and Micro Tasks. Crowdsourcing works by dismantling the fixative form of business operations while leveraging from the vast talent and skills of the crowd. Crowdsourcing attracted extensive attention from both industry and academia. Amazon Mechanical Turk (MTurk) [1] is one of the earliest crowdsourcing platforms followed by many others with different schemes to create a more balanced working environment proactively [2].

The fastest growing category of connected devices is smartphones, where users can share various types of multimedia videos, photos, and location. Every person can now participate in any crowdsourcing platform to illustrate, crowdsourcing for ride-hailing services, road conditions updates, and many other platforms. Spatial Crowdsourcing (SC), in particular, is tied to the location of the task and crowd worker. To participate in the SC task, workers share their current location with the SC platform. The SC platform calculates the distance between the task and the available crowd worker for the best match [3].

Most of the existing SC platforms assign tasks to crowd workers based on the closest available worker to the task. Upon receiving a spatial task, the SC platform assigns the first nearby worker; this approach has several weaknesses. First, it assigns tasks based on the travel distance of the worker to the task and fails to consider other upcoming tasks that might enter the system. Second, it does not take into account the location privacy-preserving [4].

This paper evaluates the performance of the proposed TASC [5], a spatial crowdsourcing task assignment approach that uses the DCentroid location privacy scheme, and the batch match approach to increase the task assignment rate and the workers' privacy [6]. The evaluation was conducted on a real crowdsourcing data set to experiment with the effectiveness of TASC. The result of the evaluation shows that the TASC scheme is positively effective and performs a preferable task assignment while preserving the location privacy of the crowd workers.

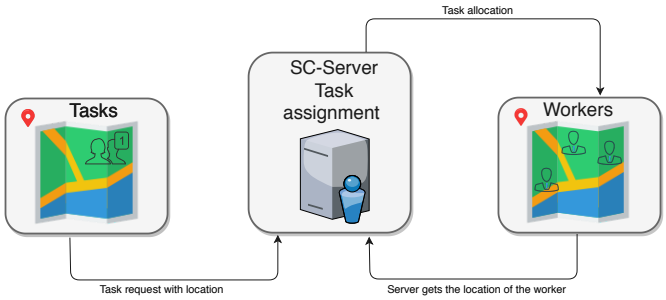


Fig. 1. Task assignment in spatial crowdsourcing

The rest of this paper is organized as follows: Section II presents the background of crowdsourcing, while Section III

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

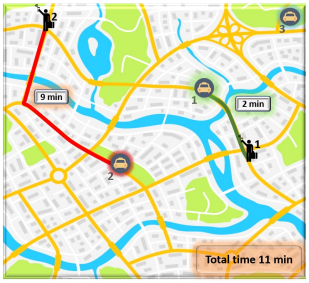


Fig. 2. Task assignment first in queue problem

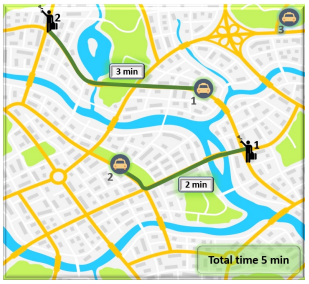


Fig. 3. TASC task assignment

introduces the system model. Section IV evaluates the system model. Finally, Section V concludes this paper and presents future work.

II. Background

A spatial crowdsourcing platform processes the location information submitted by the task requester and the worker to establishes the relationship between the task and the worker. Fig. 1 shows the workflow. In general, the SC server first collects the task information and location from the requester also the location information from the worker. The information is processed by the platform to complete the task allocation. Finally, the worker completes the spatial tasks and submits the results to the server [7].

In spatial crowdsourcing, there are three components: tasks, crowd workers, and the platform. Workers and tasks reveal their locations to the platform; then, the platform assigns workers to tasks [8]. However, the platform may disclose individual locations that present serious privacy implications such as stalking, physical surveillance, identity theft, or leakage of workers' sensitive information such as health status, relationships, and religious views. Thus, location privacy is an important aspect of SC to attract more individuals to contribute [9].

A. Task Assignment in Spatial Crowdsourcing

The current task assignment algorithms immediately dispatch the first available worker or the worker with the shortest distance or time [10]. When a request enters the system, it assigns to the first available worker or the shortest distance worker without consideration of other tasks waiting time [11]. In Fig. 2, the system immediately dispatches the first worker in the queue within the task radius; this means some crowd workers will travel a long distance to perform a task. For example, assuming w_1 and w_2 are in the system, then t_1 and t_2 enter respectively into the system, the platform will assign w_1 to t_1 and w_2 to t_2 even though w_1 might be closer to t_2. In Fig. 2, the system immediately dispatches the worker with the shortest distance; this is the optimal solution if the number of tasks is not large. This approach assigns the existing tasks in the system and does not consider the upcoming tasks, which may affect their wait-time for an assignment as Fig. 2 shown in the total task time is 11 minutes. On the other hand, Fig. 3 shows the total time of 5 minutes using TASC [12]. Spatial crowdsourcing platforms gained a lot of attention in the last decade, and the number of users is increasing. We should consider how to assign these millions of requests every day and how to save time on assigning them [13].

The researcher in [14] aimed to maximize the total score at every time instance while minimizing the travel cost of the worker. So, the study conducted three experiments to evaluate the performance of the three strategies, Which are the basic, Least Location Entropy Priority (LLEP), and Close Distance Priority (CDP) to solve the problem of maximizing the overall number of assigned tasks. Also, they demonstrated that the basic strategy used the local optimization at every time instance rather than global optimization to maximize the overall score assignment. The LLEP strategy maximizes the total score during every instance of time; thus, the total cost associated with the assigned tasks is the lowest. The CDP strategy maximizes the overall score during every time instance, while the total cost of the selected edges is the lowest. Consequently, matches with lower costs are likely to be selected first. A comparison of all these three strategies, [14] found that the basic strategy is not the best one to solve the spatial crowdsourcing problem when compared to LLEP and CDP.

B. Worker's Motivation

The worker's motivation to participate is a key feature of any crowdsourcing system and falls into two categories: self-incentivized and reward-based [15].

1) Self-incentivized: Spatial crowdsourcing workers volunteer to perform the tasks, or they have other incentives rather than receiving a reward such as documenting an event, political or religious views, and gamification [16], [17]. For example, users voluntarily report traffic information in Waze application to just get more reward points as a game.

2) Reward-based: Spatial crowdsourcing, workers will receive a specific reward for every spatial task they perform. A growing list of applications such as Uber, Left, and Fieldagent provide a price for every spatial task [9], [18], [19].

C. Location Privacy Concerns

Privacy concern is an important aspect of all spatial crowdsourcing platforms; any malicious breach to worker privacy is a threat [20]. With the Server Assigned Tasks (SAT) mode, SC-server requests the locations of tasks and workers to

perform effective task assignments; the SC-server creates the assignment links between the tasks and the workers so that they can establish a direct communication channel among themselves. Thus, the SC-server might become a single point of attack, which is a severe privacy threat [21]. The privacy concerns extend to not only the crowd workers but also to the task requesters and the SC-server [22]. The task, in some cases, contains personal information such as locations and personal interests. This information is distributed with the task; thus, the service provider may infer the requester's affiliation as an example, which should be private [23], [24].

D. DCentroid Location Privacy Scheme

Identifying the location privacy issues of crowd workers in spatial crowdsourcing has different location privacy protection techniques. However, some of the provided mechanisms techniques either are not offering acceptable results or not appropriate in every spatial crowdsourcing task [6]. For instance, [25] Pointed out that anonymization techniques cannot be applied for tasks that required crowd workers' real names.

According to [25] DCentroid is a new approach for crowd workers' location privacy in SC. This new approach is motivated by the dummy-based technique to preserve crowd workers' location privacy without significantly affecting the travel distance in task assignment as well as keeping the location hidden from the SC-Server.

Additionally, [25] the DCentroid model intents to hide the exact location of a crowd worker by sending a dummy location to the SC-Server instead of sending the real location. In this process, the crowd worker device calculates the centroid point of the dummies. It sends it to the SC-Server to calculate the estimated distance from the crowd worker to the task using the standard Euclidean distance. The SC-server assigns the task to the crowd worker based on the calculated estimated distance. So, through this process, the crowd worker's exact location information is kept hidden from the adversary [26].

Moreover, [25] an evaluation of data comparing the travel distance the non-privacy approach, that uses the exact real locations to assign the tasks, with the new approach DCentroid, showed higher travel distance error (TDE) when adding the privacy if the crowd worker is near the task [27]. However, the TDE decreased when the travel distance increased. So the proposed scheme dose not dramatically impact travel distance. Additionally, [25] the experimental results showed that the DCentroid approach is effective in crowd workers' privacy-preserving.

III. System Model

This section introduces the system model of TASC a task assignment approach with worker location privacy protection in SC. The major definitions and notations used in this framework presented in Table I.

A. TASC: Task Assignment Technique

The main idea is to inject the location privacy to the crowd workers using the DCentroid scheme p_i and defined the optimal batch size η_{batch} that can minimize the total task assignment waiting time. The mechanism of the framework calculates the travel distance $\widehat{w.l}_{i,j}$ between the available crowd workers w_i and each requested task t_i. Then, use the batch match method to select the best matching pair based on the indicated batch size η_{batch}. If there is only one task t_i in the batch, the system assigns the task t_1 to the nearest available crowd worker w_i [5].

B. Definitions

Definition 1. *Spatial Task:* Tasks with a specific locations denoted as $s.l_i$ the physically location where workers need to be to perform the task.

Definition 2. *Crowd Worker:* A crowd worker, denoted by w_i submits their location to the SC server to be assigned to nearby spatial tasks. To perform spatial tasks, crowd workers need to be online in the system.

Definition 3. *Spatial Platform:* The link between the workers and tasks. For instance, one of the core functions of any spatial platform is the assignment of tasks to suitable workers.

Definition 4. *Spatial Task Assignment:* Any spatial crowdsourcing platform set to handle a massive number of tasks and workers. For instance, assigning a large number of tasks to workers based on specific optimization objectives such as maximizing the number of assigned tasks and preventing any privacy leakage.

Definition 5. *Batch Size:* Collection of tasks in one batch η_{batch} generated by the SC server to divide the tasks into groups to make the assignments.

Definition 6. *Waiting Time:* The total waiting time η_{travel} between the requested task and the crowd worker arrival time [5].

TABLE I
Notations Definitions

Notation	Definition
w_i	Available crowd worker
t_i	Spatial task
p_i	Crowd worker pseudo location (using DCentroid scheme)
η_{batch}	Batch size
η_{travel}	Waiting time
$s.l_i$	Spatial task location
$\widehat{w.l}_{i,j}$	Estimated distance form crowd worker $_i$ to task $_j$

TASC model assigns the tasks by using the batch match method in which the CS server holds requests for a short unit of time, allowing other requests to appear in the system to create a batch of requests and find the optimal task matching. In addition, the TASC model hides the real location of a crowd worker using DCentroid which generates a dummy location the SC-server will receive, instead of sending the worker real location to increase worker privacy. The dummy location located at a certain distance from the real location as Fig. 4 explains the overall execution of the system model, which operates as follows:

1) The crowd worker device generates three dummy locations A, B, and C.

Fig. 4. System model

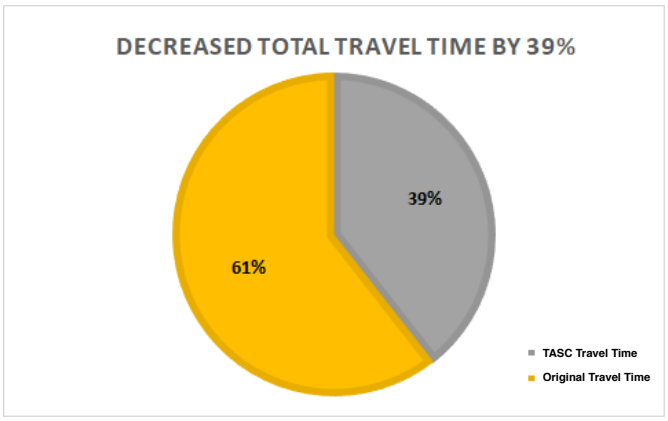


Fig. 5. Travel time

2) The crowd worker w_i device calculates the centroid point of the three dummies locations, which is not far from real location then sends it to the SC-server.
3) To assign any task, the system will wait for η_{travel} a short unit of time to generate a η_{batch} batch of requests.
4) The system will calculate the best worker and task matching if there is only one request t_1 and the time is up, the SC-server will assign the task t_i to the nearest crowd worker w_i.
5) The SC-server calculates the estimated distance between the task and the crowd worker by the standard Euclidean distance $\widehat{w.l_{i,j}}$.

$$\widehat{w.l_{i,j}} = dist(p_i, s.l_i) \quad (1)$$

where $dist$ is the Euclidean distance between two points, t_i the task location and p_i the location of the centroid point of the worker.
6) Based on the estimated distance the SC server assigns the crowd worker to the task. $\widehat{w.l_{i,j}}$.
7) Worker w_i needs to move to the location $s.l_i$ of task t_i immediately after receiving the assigned task.

IV. Performance Evaluation

This section presents the evaluation of TASC, the task assignment schemes followed by the performance, and the results.

A. Settings

Real-world dataset used to evaluate TASC. The dataset represents users' check-ins, comments, and restaurants' locations. In TASC setting, we assume that Yelp users are the crowd workers and the locations of the tasks are the restaurants [28].

Crowd worker w_i*:* The crowd workers in the dataset is the Yelp users who can write reviews and comments.

Task t_i*:* The tasks in the dataset are the restaurants with specific locations.

SC-Server *Yalp:* The platform in the dataset is Yelp application the link between the workers and tasks.

B. Evaluation and Results

We compare the performance of TASC to the current task assignment methods, which compare all workers and task pairs to task assignment and privacy. We set the number of tasks t_i as 60 and the number of workers w_i as 98. The tasks batch size η_{batch} is set to 5 tasks every η_{travel} unit of time. The crowd workers in every batch η_{batch} are sorted by the closest distance to tasks based on their dummies centroid points, then the crowd worker w_i will be assigned to the closest task. Fig. 5 shows that the performance of the TASC framework decreases the travel time for the tasks by 39%. The TASC framework demonstrates a clear advantage in terms of saved time and privacy over the current task assignment methods. To this end, the evaluation indicates that TASC travel time is less than the original travel time which is the most significant factor in SC as shown in Fig. 6 .

V. Conclusion

In this paper, we evaluated TASC a task assignment approach in spatial crowdsourcing with workers' privacy protection by using a sample of real data-set. We have proved that TASC framework shows a clear advantage in terms of saved time and privacy over the current task assignment methods. To this end, the evaluation indicates that TASC travel time is

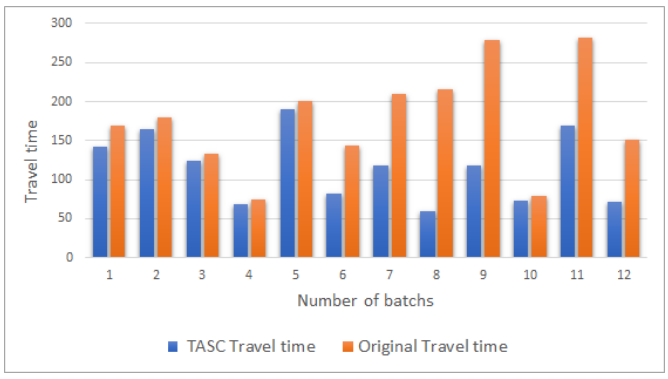


Fig. 6. TASC task assignment chart

less than the original travel time, which is the most significant factor in SC. As future work, TASC will evaluate effectiveness with quality criteria such as system overhead and the success rate of the assignment as an additional illustration of the evaluations.

REFERENCES

[1] AMT. (2005) Amazon mechanical turk. [Online]. Available: https://www.mturk.com/

[2] Y. Tong and Z. Zhou, "Dynamic task assignment in spatial crowdsourcing," *SIGSPATIAL Special*, vol. 10, no. 2, pp. 18–25, 2018.

[3] Y. Tong, Y. Zeng, Z. Zhou, L. Chen, J. Ye, and K. Xu, "A unified approach to route planning for shared mobility," *Proceedings of the VLDB Endowment*, vol. 11, no. 11, pp. 1633–1646, 2018.

[4] T. Song, Y. Tong, L. Wang, J. She, B. Yao, L. Chen, and K. Xu, "Trichromatic online matching in real-time spatial crowdsourcing," in *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*. IEEE, 2017, pp. 1009–1020.

[5] E. Aloufi, R. Alharthi, I. Alrashdi, A. Alqazzaz, D. Alsulami, and M. Zohdy, "Tasc: Efficient task assignment in spatial crowdsourcing with workers privacy protection," in *2020 IEEE International Conference on Electro Information Technology (EIT)*. IEEE, 2020.

[6] R. Alharthi, E. Aloufi, A. Alqazzaz, I. Alrashdi, M. Zohdy, and J. Rrushi, "Protecting location privacy for crowd workers in spatial crowdsourcing using a novel dummy-based mechanism," *IEEE Access*, vol. 8, no. 1, pp. 114 608–114 622, 2020.

[7] Y. Liu, B. Guo, Y. Wang, W. Wu, Z. Yu, and D. Zhang, "Taskme: Multi-task allocation in mobile crowd sensing," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, pp. 403–414.

[8] H.-F. Ting and X. Xiang, "Near optimal algorithms for online maximum edge-weighted b-matching and two-sided vertex-weighted b-matching," *Theoretical Computer Science*, vol. 607, pp. 247–256, 2015.

[9] H. To, G. Ghinita, and C. Shahabi, "A framework for protecting worker location privacy in spatial crowdsourcing," *Proceedings of the VLDB Endowment*, vol. 7, no. 10, pp. 919–930, 2014.

[10] H. To, L. Fan, L. Tran, and C. Shahabi, "Real-time task assignment in hyperlocal spatial crowdsourcing under budget constraints," in *2016 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2016, pp. 1–8.

[11] B. Guo, Y. Liu, W. Wu, Z. Yu, and Q. Han, "Activecrowd: A framework for optimized multitask allocation in mobile crowdsensing systems," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 3, pp. 392–403, 2016.

[12] A. S. Fonteles, S. Bouveret, and J. Gensel, "Towards matching improvement between spatio-temporal tasks and workers in mobile crowdsourcing market systems," in *Proceedings of the third acm sigspatial international workshop on mobile geographic information systems*, 2014, pp. 43–50.

[13] L. Tran, H. To, L. Fan, and C. Shahabi, "A real-time framework for task assignment in hyperlocal spatial crowdsourcing," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 9, no. 3, pp. 1–26, 2018.

[14] H. To, C. Shahabi, and L. Kazemi, "A server-assigned spatial crowdsourcing framework," *ACM Transactions on Spatial Algorithms and Systems (TSAS)*, vol. 1, no. 1, pp. 1–28, 2015.

[15] L. Zhang, X. Lu, P. Xiong, and T. Zhu, "A differentially private method for reward-based spatial crowdsourcing," in *International Conference on Applications and Techniques in Information Security*. Springer, 2015, pp. 153–164.

[16] H. Xiong, D. Zhang, G. Chen, L. Wang, and V. Gauthier, "Crowdtasker: Maximizing coverage quality in piggyback crowdsensing under budget constraint," in *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2015, pp. 55–62.

[17] T. Yan, V. Kumar, and D. Ganesan, "Crowdsearch: exploiting crowds for accurate real-time image search on mobile phones," in *Proceedings of the 8th international conference on Mobile systems, applications, and services*, 2010, pp. 77–90.

[18] Uber. (2009) Uber. [Online]. Available: https://www.uber.com/

[19] Waze. (2007) Waze. [Online]. Available: https://www.waze.com/

[20] K. Yang, K. Zhang, J. Ren, and X. Shen, "Security and privacy in mobile crowdsourcing networks: challenges and opportunities," *IEEE Communications Magazine*, vol. 53, no. 8, pp. 75–81, 2015.

[21] H. To and C. Shahabi, "Location privacy in spatial crowdsourcing," in *Handbook of Mobile Data Privacy*. Springer, 2018, pp. 167–194.

[22] H. To, G. Ghinita, L. Fan, and C. Shahabi, "Differentially private location protection for worker datasets in spatial crowdsourcing," *IEEE Transactions on Mobile Computing*, vol. 16, no. 4, pp. 934–949, 2017.

[23] B. Zhu, S. Zhu, X. Liu, Y. Zhong, and H. Wu, "A novel location privacy preserving scheme for spatial crowdsourcing," in *2016 6th International Conference on Electronics Information and Emergency Communication (ICEIEC)*. IEEE, 2016, pp. 34–37.

[24] Y. Gong, C. Zhang, Y. Fang, and J. Sun, "Protecting location privacy for task allocation in ad hoc mobile cloud computing," *IEEE Transactions on Emerging Topics in Computing*, vol. 6, no. 1, pp. 110–121, 2015.

[25] R. Alharthi, E. Aloufi, A. Alqazzaz, I. Alrashdi, and M. Zohdy, "Dcentroid: Location privacy-preserving scheme in spatial crowdsourcing," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*. IEEE, 2019, pp. 0715–0720.

[26] A. Liu, Z.-X. Li, G.-F. Liu, K. Zheng, M. Zhang, Q. Li, and X. Zhang, "Privacy-preserving task assignment in spatial crowdsourcing," 2017.

[27] Y. Shen, L. Huang, L. Li, X. Lu, S. Wang, and W. Yang, "Towards preserving worker location privacy in spatial crowdsourcing," in *2015 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2015, pp. 1–6.

[28] yelp. (2004) yelp. [Online]. Available: https://www.yelp.com/

Implementation on Design of a Novel Frequency Selective Surface

Kousik Roy
Department of ECE
Asansol Engineering College
Asansol, West Bengal, India
kousikroy002@gmail.com

Pratik Dey
Department of ECE
MAKAUT
Haringhata, West Bengal, India
pratikdey87@gmail.com

Naimul Hasan
Department of ECE
SETGOI
Durgapur, West Bengal, India
naimulhasan1984@gmail.com

Bhaskar Roy
Department of AEIE
Asansol Engineering College
Asansol, West Bengal, India
papan.ec11@gmail.com

Susanta Kumar Parui
Department of ECE
IIEST
Shibpur, West Bengal, India
arkapv@yahoo.com

Abstract— **This paper deals with the design of a compact planar passband frequency selective surface (FSS) operating at the resonance frequency of 20 GHz. A sole layer of the cross type with the interdigital capacitance connecting the corners of metallic patches printed on single side of the substrate is used in this FSS structure. Here the transmission and reflection coefficient of the projected FSS gives an accurate resonance constancy for both of polarization and various incident angles. A unit cell which is mainly a cross shaped is designed here and here the performance of this unit cell at a definite frequency range is analyzed. The S Parameters which are S_{11} and S_{21} for TE_{00} mode that is for Perpendicular Polarization is shown in this paper. Along with that the S Parameters which are S_{12} and S_{22} for TM_{00} mode that is for Parallel Polarization is also pointed in this paper.**

Keywords- FSS, Polarization, unit cell, periodic structure

I. INTRODUCTION

Mainly a Frequency Selective Surface (FSS) is a flat structure of composite materials which is aimed to be transparent in some frequency bands while reflective or absorbing at other frequencies. FSS is a collection of metal screens which is normally very smooth. A two-dimensional periodic element mainly constructs the structure of an FSS which is clearly drawn in the figure and it provides the frequency filtering behavior which is very near to those of frequency filters used in Radio Frequency (RF) circuits. Different geometrical elements which are mainly of shaped of a dipole, square patch, circular patch, cross, and Jerusalem cross are focused in this FSS structure, for which the whole domain functions have been successfully attended [1-4]. The widely considered FSS structure are Circular rings as well as square loops [5-7]. These geometrical structures are pointed out in Figure 1.

FSS are basically a diffraction grating which will diffract an applied wave into different directions. Figure 2 depicts the Grating lobes of this unit cell FSS. A tremendous amount of control over the form of the response of an FSS can be recognized by using a multilayered resonance structures which can join longitudinal resonance, absorption, transverse resonance and diffraction into a particular device. FSS build with multilayer structures are better reactive to incident angle. If a cascaded periodic structure is used in the formation of this FSS then we obtain a broader top and quicker roll off. The bandwidth is also changed at a considerable manner with the change of incident angle. The FSS designing and investigation along with their parameter's calculation is a multidisciplinary and exciting details of antenna as well as electromagnetic theory. If we change a parameter in an FSS then this change has some electromagnetic significances and subsequently it effects the frequency filtering also.

978-1-7281-9615-2/20/$31.00©2020 IEEE

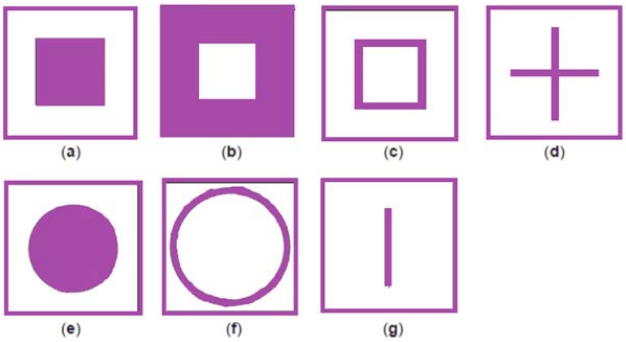


Figure 1. Some typical FSS element shapes

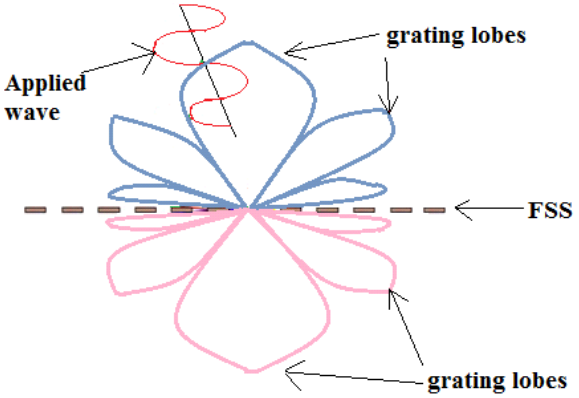


Figure 2. Grating lobes of this unit cell FSS

II. FSS FUNDAMENTALS

A Frequency Selective Surface (FSS) is recognized by a periodic array containing conducting patch or aperture elements. The array element sort which is mainly patch or aperture, can determines the low pass or high pass spectral activities possessed by an FSS, which is as similar frequency filters in Radio Frequency (RF) [8]. Figure 3 shows a two-dimensional planer periodic structure.

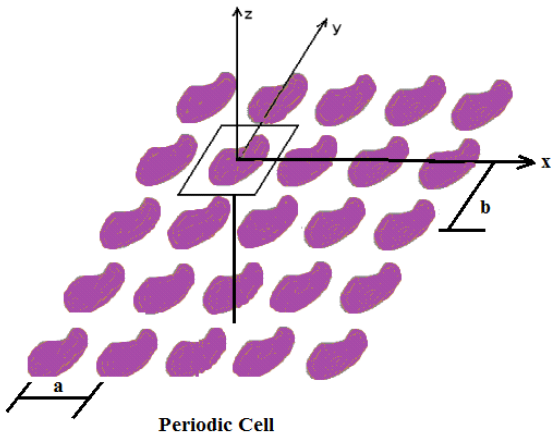


Figure 3. Geometry of a two-dimensional periodic structure

Numerous design restrictions and values are interconnected with the periodic structure such as element shape, size, lattice geometry, dielectrics, grating lobes and Wood's variance. Basically, the circular, rectangular, dipole, cross dipole, Jerusalem cross, tripole, three or four-legged dipole, ring, square loop and gridded square loop are the different fundamental shapes used in an FSS [8]. The different component formats which are developed by join up or changing of plain elements which are shown above are used and the elements of very complex shapes for example convolutional or fractals are also used [9-10]. Figure 4 depicts the FSS defined as a periodic array of aperture or patch elements. The FSS which are mainly of aperture component

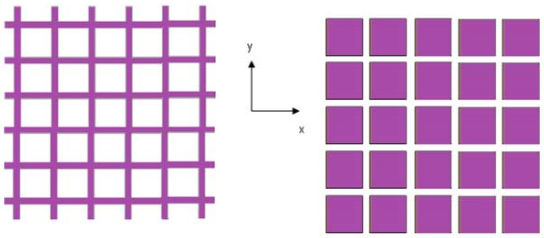


Figure 4. FSS Elements (a) type Aperture and (b) type Patch

can reproduces at low frequencies and transmits at high frequencies (nearer to high pass filter), on the other hand patch type FSS can reproduces at high frequencies and transmits at low frequencies (nearer to low pass filter) [9]. FSS analysis is done by various techniques. The equivalent circuit representation is one of the simplest techniques. An isolated patch element in a periodic array formed by various strip elements which forms as inductive and capacitive device on a transmission line are discussed in this study. The reflection and transmission coefficients of the FSS are recognized by the description of this circuit.

III. DESIGN AND SIMULATION

The FSS seeing as an infinite structure is analyzed here by us in ANSYS HFSS. That means this FSS will be irradiated by a pure TEM mode. Some transmitted waves can illuminate that FSS when we use this FSS in our particular application and it is of a pure TEM mode or TEM waves which are coming from a far distance and it is of a plane wave nature. So, it is needed to generate that certain situation by a software. This FSS are used in some antennas or some far distant object. Incident waves coming from a distant object and it is a purely TEM modes. This situation is generated by a software using Floquet Port and it is analyzed by Floquet Mode analysis. Both the reflection and the transmission behavior of our unit cell are changed by the angle of incidence of this incident waves. The design of our proposed unit cell FSS is shown in Figure 5. It is a square substrate of 12.5 mm length from which a slot of cross shape is removed. The slot dimension is plotted in the figure 5. Now in figure 6 and figure 7 we have given the design of Latice Pair 1 and Latice Pair 2 respectively. And in figure 8 and figure 9 we

have given the design of Floquet Port 1 and Floquet Port 2 respectively.

Figure 10 depicts the S Parameters which are S_{11} and S_{21} for TE $_{00}$ mode that is for Perpendicular Polarization. Now S_{11} is for Reflection Characteristics and S_{21} is for Transmission Characteristics. So, from figure 10 we see that this unit cell FSS is transparent at our operating frequency 20 GHz as it has a reflection of only -20.144 dB which is near about -20 dB that is only 2 % of reflection. Similarly figure 11 depicts the S Parameters which are S_{12} and S_{22} for TM_{00} mode that is for Parallel Polarization. Now S_{12} is for Reflection Characteristics and S_{22} is for Transmission Characteristics for the TM_{00} modes or Parallel Polarization. So, from Figure 10 and Figure 11 it is cleared that both these reflection and transmission characteristics are same for these two modes (TE_{00} and TM_{00}). So, it is cleared that this unit cell FSS is working very well for both the Parallel as well as Perpendicular Polarization. Now Figure 12 depicts the VSWR result for this unit cell FSS at TE 00 mode. From figure 12 it is cleared that VSWR is approx. 1 at our operating frequency 16.5 GHz. So, it is a very good condition for the working of this unit cell FSS. Figure 13 depicts the Transmission Characteristics for both the Perpendicular Polarization (TE_{00} mode) and Parallel Polarization (TM_{00} mode). From this figure it is cleared that this unit cell FSS has equally transmitted for both Parallel Polarization and Perpendicular Polarization signal. Figure 14 depicts the Transmission Characteristics for Perpendicular Polarization Higher Order Mode (TE_{10}). Figure 15 depicts the Electric Field Vector at unit cell FSS. From this figure it is cleared that Electric Field is perpendicular to this unit cell.

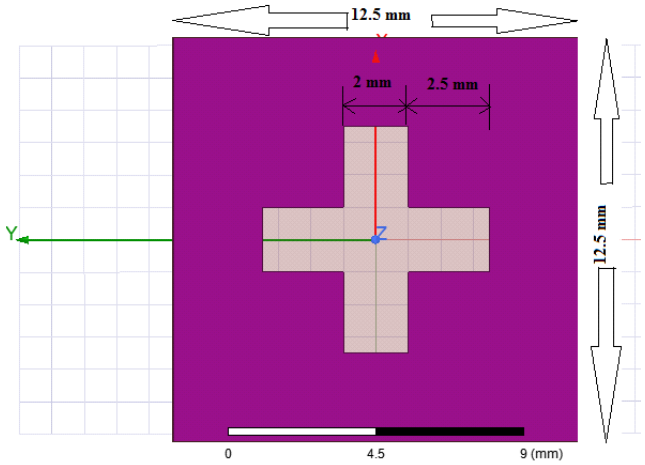


Figure 5. Design of our unit cell FSS

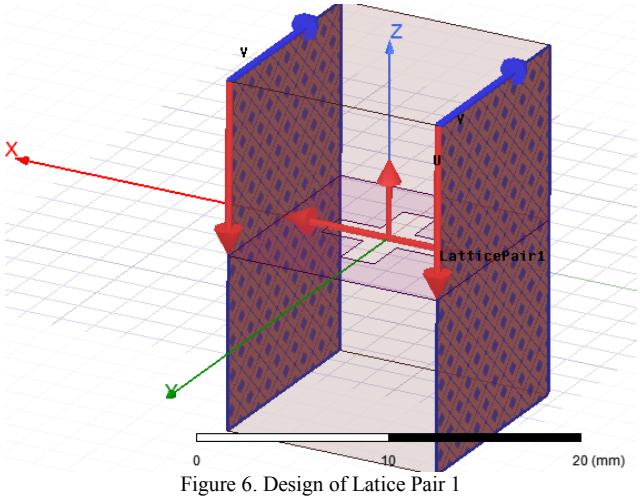


Figure 6. Design of Latice Pair 1

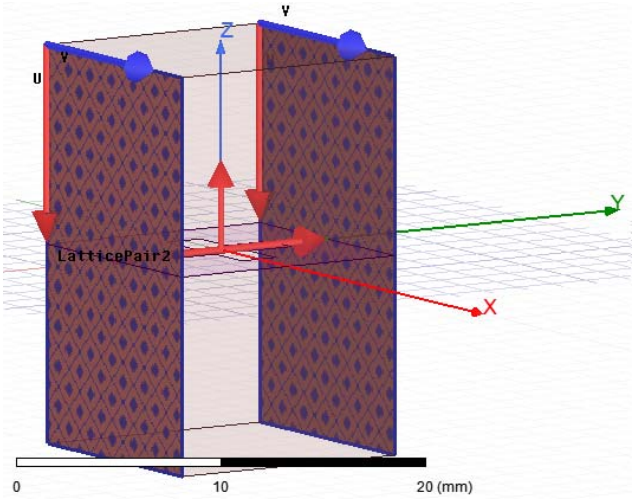


Figure 7. Design of Latice Pair 2

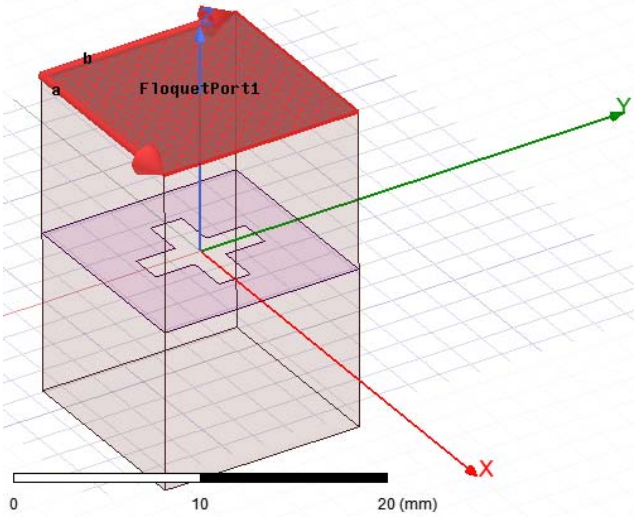


Figure 8. Design of Floquet Port 1

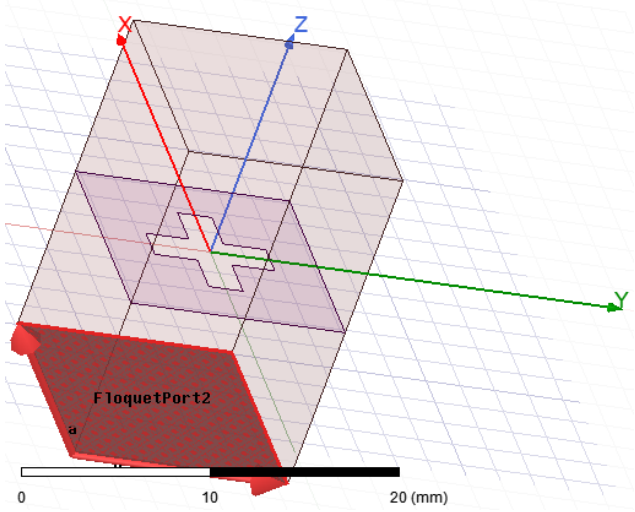


Figure 9. Design of Floquet Port 2

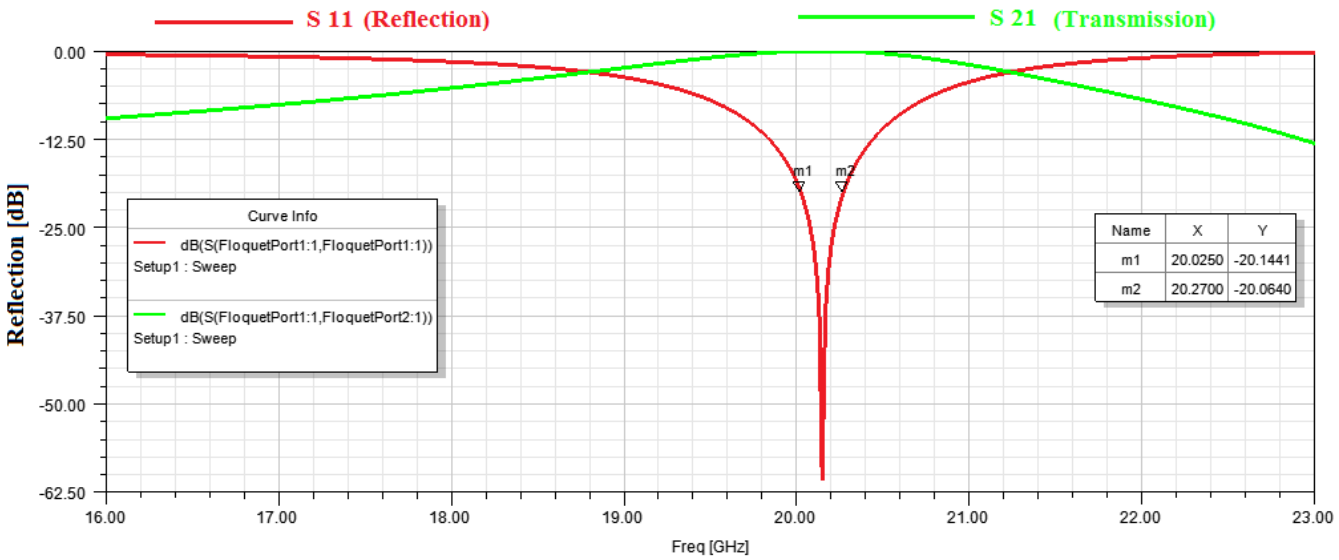


Figure 10. S Parameter Result (S_{11} and S_{21}) for TE_{00} Mode

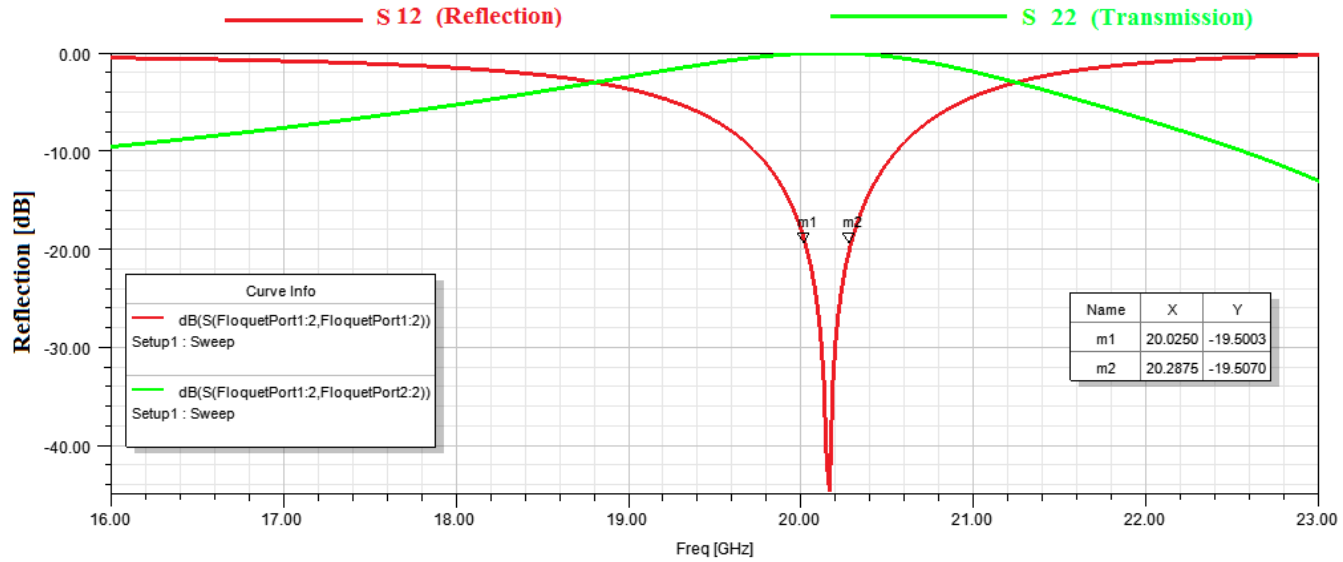


Figure 11. S Parameter Result (S_{12} and S_{22}) for TM_{00} Mode

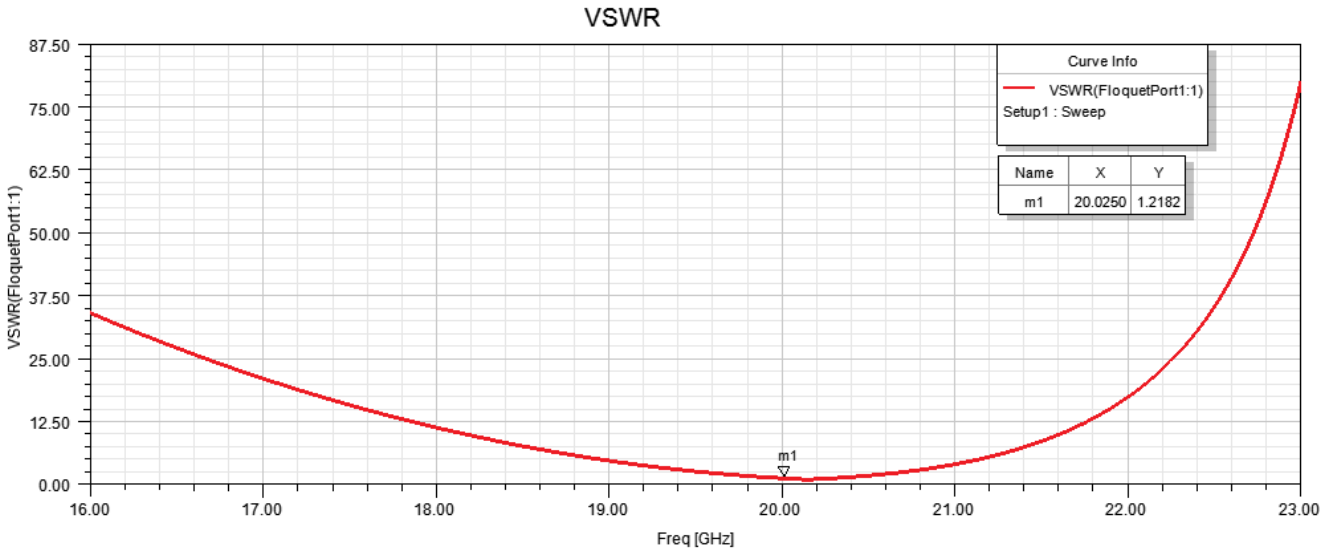


Figure 12. Variation of VSWR with respect to Frequency

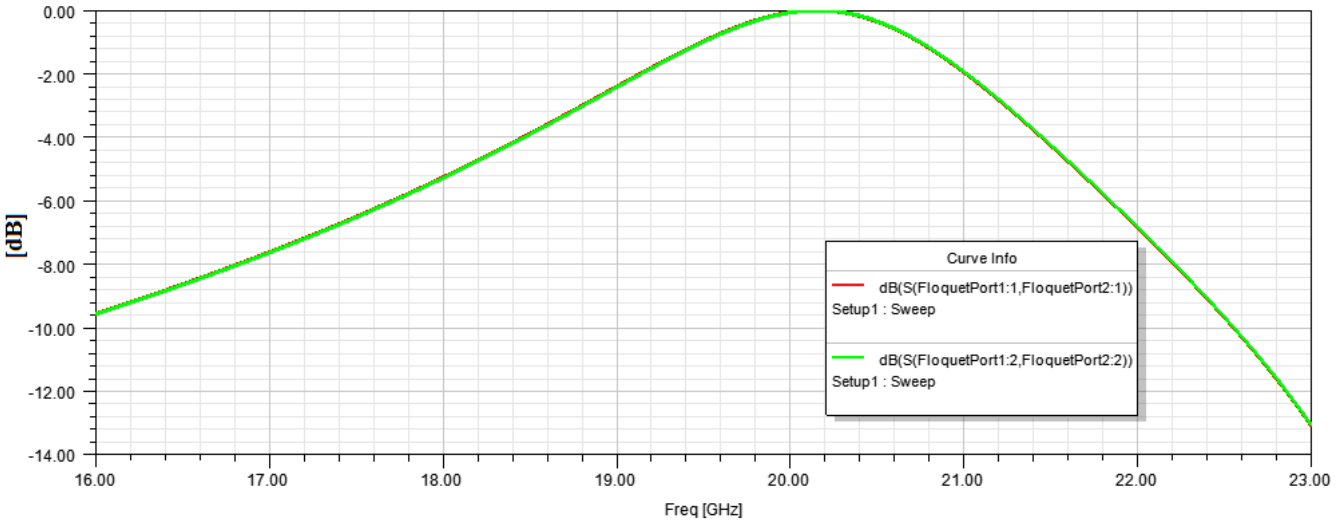


Figure 13. Transmission Characteristics for both the Perpendicular Polarization (TE_{00}) and Parallel Polarization (TM_{00})

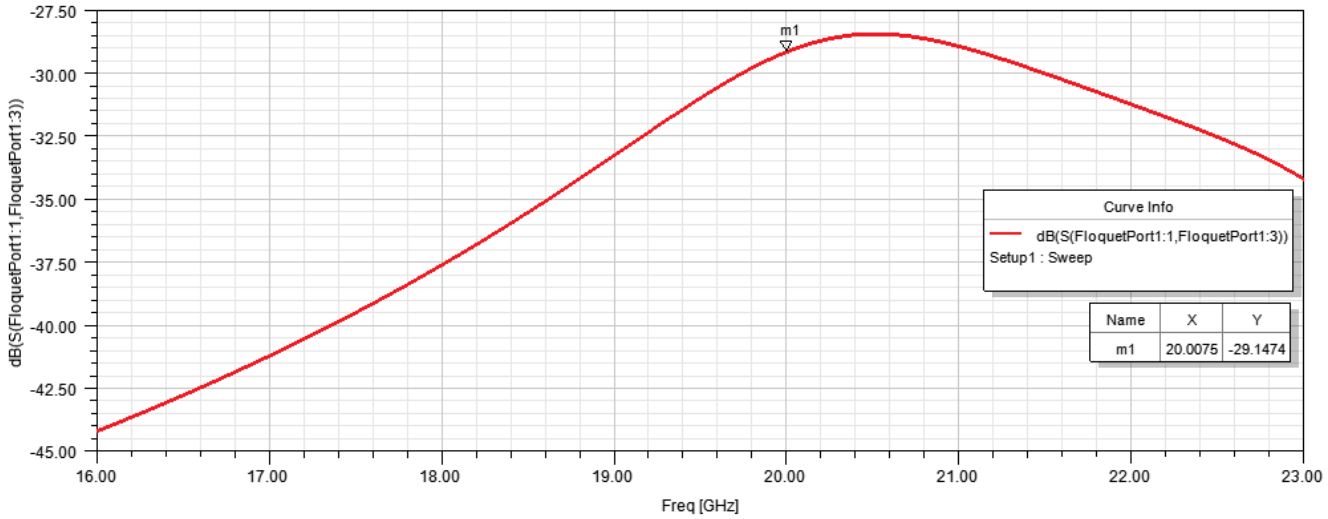


Figure 14. Transmission Characteristics for Perpendicular Polarization Higher Order Mode (TE_{-10})

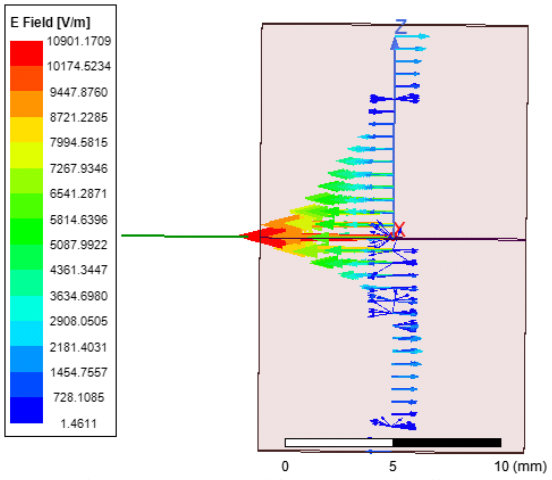


Figure 15. Electric Field Vector at unit cell FSS

IV. CONCLUSION

A compact frequency selective surface with passband properties at the resonance frequency of 20 GHz is provided in this paper. Resonance constancy at an incident angle with independent polarization as well as its compactness have disclosed by this proposed FSS. This FSS structure is well sutaibly working for both the perpendicular polarization as well as parallel polrization.So any electromagnetic wave acting as a vertical polarization or horizontal polarization is perfectly frequency filtered by this FSS. A nearly linear phase deviation over the frequency band which makes it supportive for accumulative radiation characteristics for example gain, directivity, front to back ratio of the antennas with omni-directional or bi-directional radiation (slot antennas as well as CPW antenna) over a wide bandwidth when utilized as a reflector is provided by this projected FSS.

REFERENCES

[1] C. H. Tsao and R. Mittra, "Spectral-domain analysis of frequency selective surfaces comprised of periodic arrays of cross dipoles and Jerusalem crosses," IEEE Trans. Antennas Propagat., vol. AP-32, no. 5, pp. 478-486, May 1984.

[2] B. Li and Z. X. Shen, “Three dimensional band-pass frequency selective structures with multiple transmission zeros,” IEEE Trans. Microw. Theory Tech. 61, 3578–3589 ,2013.

[3] B. A. Munk, Frequency Selective Surfaces: Theory and Design (Wiley, New York, 2000).

[4] M. Al-Joumayly and N. Behdad, “A new technique for design of low-profile, second-order, bandpass frequency selective surfaces”, IEEE Trans. Antennas Propag. 57, 452–459 ,2009.

[5] C. H. Tsao, "Spectral-domain approach for analyzing scattering from frequency selective surface," Ph.D. Dissertation, University of Illinois, Urbana, IL, 1981.

[6] G. Q. Luo, W. Hong, Z. C. Hao, B. Liu, W. D. Li, J. X. Chen, H. X. Zhou, and K. Wu, “Theory and experiment of novel frequency selective surface based on substrate integrated waveguide technology,” IEEE Trans. Antennas Propag. 53, 4035–4043 2005.

[7] X. Sheng, J. Fan, N. Liu, and C. Zhang, “A miniaturized dual-band FSS with controllable frequency resonances,” IEEE Microwave and Wireless Components Lett., vol. 27, no. 10, pp.915-917, Oct. 2017.

[8] T. K. Wu, Frequency Selective Surface and Grid Array, New York, NY: John Wiley & Sons, 1995.

[9] M. Yan, J. Wang, and H. Ma, “A tri-band, highly selective, bandpass FSS using cascaded multilayer loop arrays,” *IEEE Trans. Antennas and Propag.*, vol. 64, no. 5, pp. 2046-2049, May 2016.

[10] M. Z. Joozdani and M. K. Amirhosseini, “Equivalent circuit model for the frequency selective surface embedded in a layer with constant conductivity,” IEEE Trans. Antennas and Propag., vol. 65, no. 2, pp. 705-712,Feb.2017.

Wideband Dual-Cut Circular Ring based Linear-Cross and Linear-Circular Polarizing Reflector

Mohammad Abdul Shukoor
Electrical engineering
IIT Palakkad
Kerala, India
abdulshukoor872@gmail.com

Sukomal Dey
Electrical engineering
IIT Palakkad
Kerala, India
sdey28@iitpkd.ac.in

***Abstract*— In this paper, a reflective type wideband polarization converter is proposed. It can convert both linear to linear and linear to circular for different frequency bands. It consists of a simple dual-cut circular split ring resonator, printed on a thin FR-4 dielectric substrate terminated with the ground. Simulation result shows, an *x/y* polarized incident wave conversion to *y/x* reflected wave in three bands (4.49-4.63 GHz, 7.11-10.75 GHz, 14.94-16.93 GHz) and *x/y* polarized incident wave returns the circularly polarized reflected wave in three different bands (4.86-5 GHz, 6.09-6.46 GHz, 11.83-14.03 GHz). The proposed structure is compact with the unit-cell periodicity of 0.14λ and overall thickness 0.047λ, where λ corresponds to the lowest frequency of the operational band. The performance of the proposed structure is investigated under different oblique incidences. Surface current patterns are studied to analyze the reason behind the wideband response. Authors strongly believe that the proposed structure has the potential to be used in different applications where both linear-cross and linear-circular conversion are the primary concerns.**

Keywords*— Wideband, Linear-cross, Linear-Circular, Frequency Selective Surface (FSS), Polarization Conversion Ratio (PCR).

I. INTRODUCTION

The orientation of the electric field oscillations with respect to time in a plane perpendicular to the direction of propagation in free-space is referred to as Electromagnetic (EM) wave polarization. Controlling and manipulating the states of EM wave polarization has significance in many practical applications like RCS reduction [1], Quarter and halfwave plates [2], and mutual coupling reduction [3]. Conventionally, the polarization conversion can be achieved with the structural anisotropy available in nature available crystals called Birefringence effect [4]. Due to weak anisotropy, these structure needs larger thickness compared to operating wavelength for the wave to undergo considerable phase shift.

On the other hand, Metamaterials have recently gained interest in many applications such as negative refraction [5], invisible cloaking [6], and reverse doppler effect [7], due to their unusual exotic behaviour not commonly found in natural materials. Multiple polarization converters based on artificial anisotropy and chirality have been reported in the literature [8]. However, the narrow bandwidth nature of these crystals is limiting its usage in practical applications. Later, stacking of different layers [9], gold-helix structures [10] and anisotropic high impedance surfaces (HIS) [11] are also used for the extension of bandwidth. Due to the unacceptable thickness of these devices, it is hard to interface with complex sub-miniaturized systems. 2D metamaterial-based plane polarizer is a promising solution in terms of bandwidth and physical thickness. Planar polarizers like double V-shaped patches[12], L-shaped patches [13], square split ring resonator [14], and

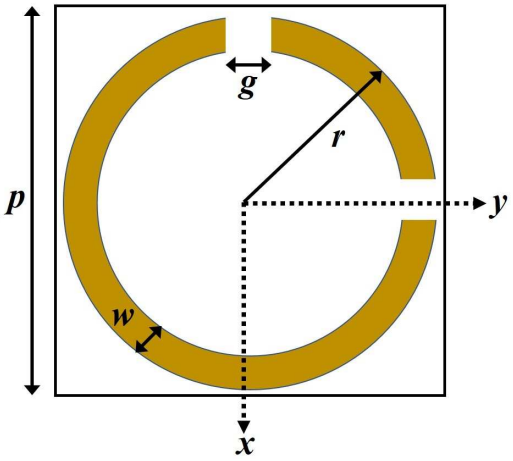


Fig. 1 Unitcell of the proposed polarization converter.

oval ring resonator [15] are proposed in literature to attain linear to cross-polarization conversion over the broadband. A broadband linear-cross polarizer based on concentric circular-cut rings connected to single dipole tilted at 45° degree with the incident wave is illustrated in [16], maintaining 90% PCR from 8.8-22.4 GHz with 86% fractional bandwidth (FBW). Mourad *et al.* proposed a reflective type linear-circular polarization converter with symmetric anisotropic unitcell exhibiting the axial ratio less than 3dB over 6.9 GHz to 14.1 GHz (66.6% FBW) [17]. A simple I-shaped linear to circular polarization reflector is proposed with 3 dB axial ratio bandwidth of 11.65 GHz (from 9.45 to 21.1 GHz) in [18]. Besides, few research works have been conducted that simultaneously transform linear polarization into cross-polarization and circular polarization over wideband.

Multi-polarization conversion is also achieved in different frequency bands using double L-shaped metasurface in [19]. Zheng *et al.* demonstrated a wideband linear to cross and linear to circular reflective polarizer in [20]. This design consists of two meander lines and a cut dipole printed on the grounded dielectric substrate, which exhibits linear to cross conversion in the lower frequency band (6.53 GHz to 12.07 GHz) with 59.6% FBW and linear to circular conversion in the higher frequency band (6.53 GHz to 12.07 GHz) with 13% FBW.

In this paper, the response of a simple dual-cut circular ring based reflective type polarization converter is demonstrated, which shows both linear to linear, and linear to circular polarization conversion simultaneously in multiple frequency bands. The top FSS is printed on a grounded substrate to make the transmission zero. Since the reflective type structures are inherently broadband, this topology is

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

adopted here. The cause of polarization conversion has been studied.

This paper is organized as follows. In section II, the designed parameters of unitcell are discussed with simulation results. The reason behind the different polarization mechanism is detailed in section III. The analysis of surface current patterns of the top metasurface and ground are illustrated in section IV. Utility of the proposed design for RCS reduction is demonstrated in section V. Finally, the conclusion has been drawn in section VI.

II. Design of the Unitcell

The proposed reflective type broadband polarization converter is composed of a circular ring resonator with two identical cuts at 90° apart, as shown in Fig. 1. It is printed on the top of a commercially available 3.2 mm thin FR-4 dielectric substrate having relative permittivity $\varepsilon_r = 4.4$, and loss tangent $tan\,\delta = 0.02$. The other side of the substrate is terminated with the ground. A thin copper film with conductivity $\sigma = 5.8 \times 10^7$ S/m and thickness t = 0.035 mm is used as the metallic layer for both top FSS and ground. The analysis of the unitcell has been carried out by a Finite Element Method (FEM) based EM solver ANSYS HFSS. Master-slave periodic boundary conditions were applied along the *x*-*y* plane and the incident wave along the *z*-direction. The performance of the unitcell is optimised and its dimensions are: p = 9.4 mm, r = 4 mm, w = 0.4 mm, and g = 0.5 mm. Here, p is the periodicity of the unit-cell, r is the radius of the ring, w is the width of the ring, g is the width of the two similar gaps, as shown in Fig. 1.

III. Results and Discussions

Consider, a *y*-polarized EM wave incident on the metasurface. The co-pol reflectance ($r_{yy} = |\overrightarrow{E_r^y}|/|\overrightarrow{E_i^y}|$), and cross-pol reflectance ($r_{xy} = |\overrightarrow{E_r^x}|/|\overrightarrow{E_i^y}|$) are defined. Here, $\overrightarrow{E_i}$ and $\overrightarrow{E_r}$ represents the electric field intensities of the incident and reflected waves. The efficiency of the linear to cross (*y to x*) conversion mechanism is determined by Polarization Conversion Ratio (PCR), and is defined as $PCR = r_{xy}^2/(r_{xy}^2 + r_{yy}^2)$. The simulated r_{yy} and r_{xy} are shown in Fig. 2, PCR and ECR in Fig. 3. From Fig. 2, it seems that there are a total of five co-pol reflection dips with near 0 dB cross-pol reflection coefficient. At these resonances, the perfect *y*-*x* conversion would be done, as shown in Fig. 3. At frequencies 4.5 GHz, 7.6 GHz, 10.1 GHz, 15.5 GHz, and 16.3 GHz the observed PCRs are 97.27%, 99.79%, 99.84%, 99.3% and 99.27%, respectively. The PCR is greater than 90% for two wide bands from 7.11-10.75 GHz (3.64 GHz bandwidth), and 14.94-16.93 GHz (1.99 GHz bandwidth). The denominator part of PCR is defined as Energy Conversion Ratio (ECR), $ECR = r_{xy}^2 + r_{yy}^2$, which illustrates energy loss experienced by the wave in the process of polarization conversion (major contribution is due to the dielectric losses and metallic losses). From the response in Fig. 3, ECR is almost 90% up to 14 GHz, and then it approaches 60 % at 18 GHz. The PCR is nearly 50% in two wide bands (5-7 GHz and 11-13 GHz range), which means that 50% of energy only is converted into its orthogonal direction.

In the next section, the azimuth angle (η) is calculated from the expression given

$$\eta = tan^{-1}\left(|\overrightarrow{E_r^x}|/|\overrightarrow{E_r^y}|\right) \quad (1)$$

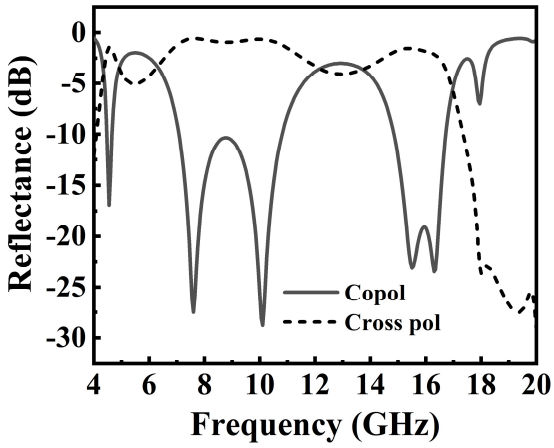


Fig. 2. Co-pol and Cross-pol reflectance of the proposed converter.

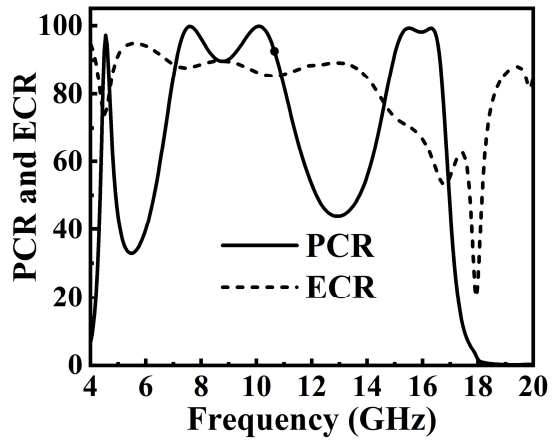


Fig. 3. PCR and ECR response of the proposed converter.

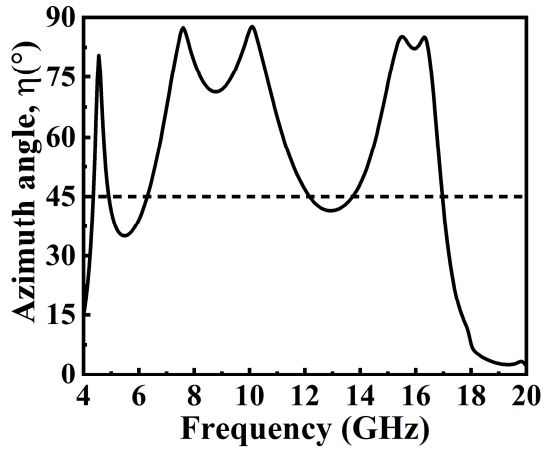


Fig. 4. Azimuth angle of the proposed converter.

The simulated response of the azimuth angle is depicted in Fig. 4. Azimuth defines the angle made by reflected wave with major- axis. From Fig. 4, at resonant peaks, the azimuth angle approaches 90°, illustrating the ideal orthogonal conversion of the reflected wave with respect to the incident wave. It is observed that the azimuth angle lies near to 45° ± 5°in three frequency bands, which corresponds to the circularly polarized reflection. The axial ratio has been calculated to

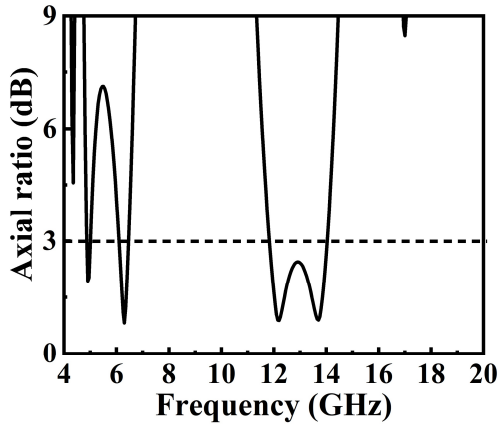


Fig. 5. Axial ratio variation of the proposed converter.

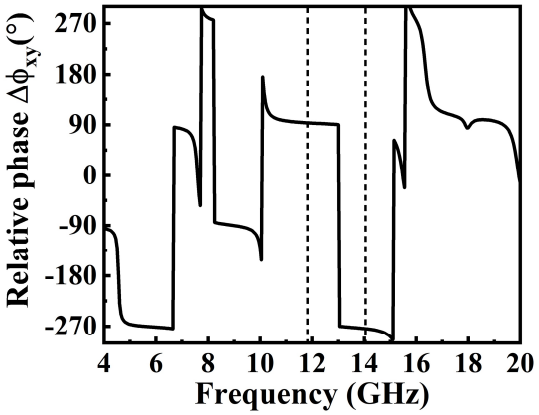
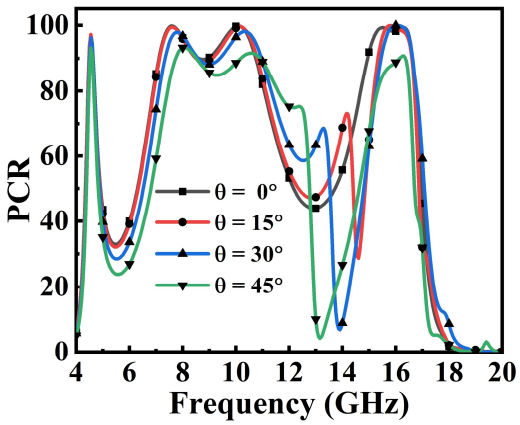


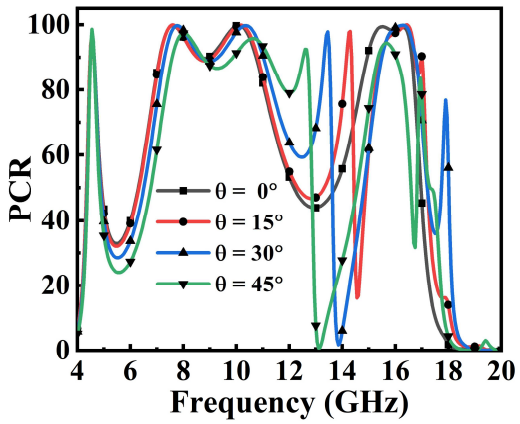
Fig. 6. Relative phase of reflected components.

determine the strength of circular polarization, from the expression given in [20], as shown in Fig.5. The axial ratio is less than 3 dB for three frequency bands, i.e. 4.86-5 GHz, 6.09-6.46 GHz, 11.83-14.03 GHz. At these band, the co-pol and cross-pol reflectances are equal and azimuth angle near to 45° satisfying the conditions for circular polarization. However, from these, the polarity of the circularly polarized wave couldn't be justified. The relative phase between co-pol and cross-pol reflectance is determined [20], and its response is depicted in Fig. 6. Here, it is clear that from 11.83 GHz to 13.03 GHz, $\Delta\phi_{xy} = +90°$ and from 13.03 GHz to 14.03 GHz, $\Delta\phi_{xy} = -270°$ (= +90°), which implies the reflected wave is Left Circularly Polarized (LCP) for the *y*-polarized incidence.

In this section, the variation in the performance of the proposed metasurface under different oblique incidences for Transverse Electric (TE) and Transverse Magnetic (TM) modes is analyzed, the PCR and Axial ratio variations for different angles are plotted, as shown in Fig. 7 and Fig. 8, respectively. From Fig. 7, it is observed that the proposed design is showing stable 90% PCR response up to 45° for both TE and TM incidences with the negligible bandwidth reduction. But it shows considerable disturbance in between the bands (circular polarization section) for both TE and TM. The dependence of the axial ratio for oblique TE incidence is studied, which depicts that the axial ratio is more sensitive with the incident angle (see Fig. 8). From the above response, this design is stable up to 45° for linear-cross conversion and up to 30° for linear-circular conversion.



(a)



(b)

Fig. 7. Variation in the PCR under different oblique incidence (a) TE mode. (b). TM mode.

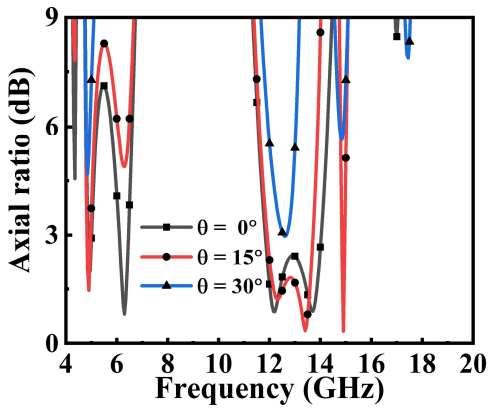


Fig. 8. Absorptivity response for different polarization angle under normal incidence.

IV. Surface Current Distributions

The reason behind the wideband performance of reflective type polarization converter is illustrated in this section. For this purpose, the surface current patterns of the top FSS and ground at different frequencies are illustrated in Fig. 9. At 4.55 GHz (see Fig. 9(a)), the surface current on the top FSS is concentrated at the middle of the larger arc, and it is anti-parallel with the currents in ground plane (see Fig. 9(b)).

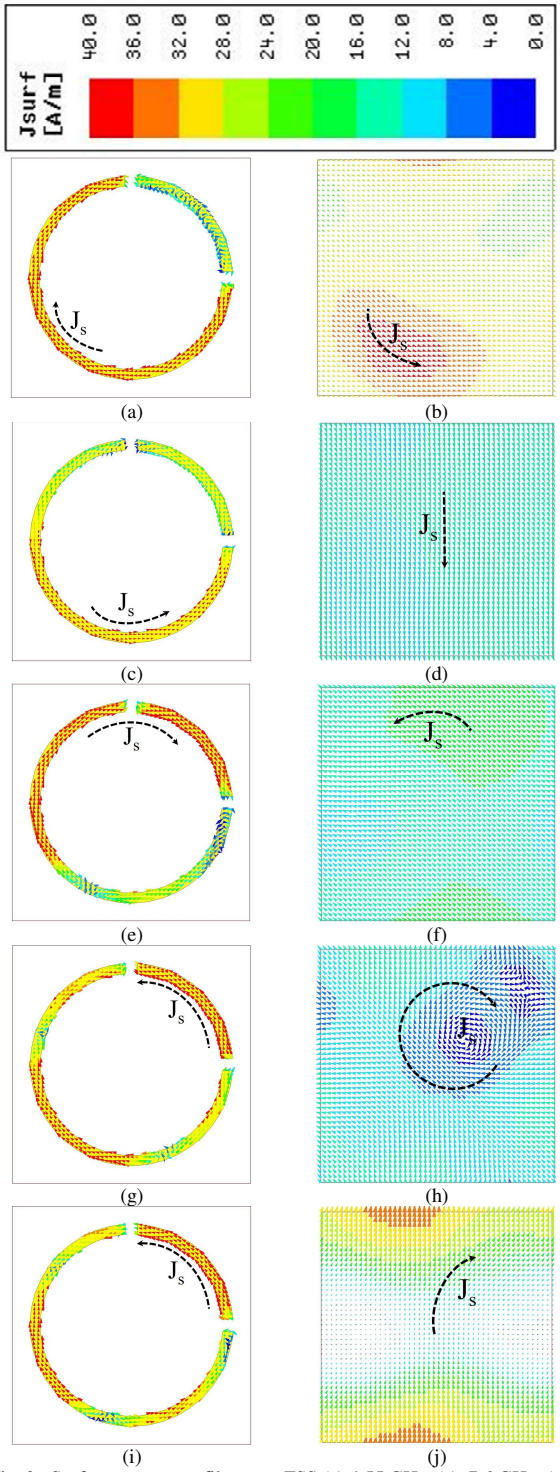


Fig. 9. Surface current profile at top FSS (a) 4.55 GHz. (c). 7.6 GHz. (e). 10.1 GHz. (g). 15.5 GHz. (i). 16.3 GHz; at Ground (b) 4.55 GHz. (d). 7.6 GHz. (f). 10.1 GHz. (h). 15.5 GHz. (j). 16.3 GHz;

These anti-parallel currents constitute a magnetic dipole resembling magnetic or dielectric resonance. The same explanation is applicable at different regions of the FSS for different frequencies, i.e. at 10.1 GHz, 15.5 GHz, 16.3 GHz. From the response at 15.5 GHz (see Fig. 9(h)), it seems clear that the surface currents in the ground plane are forming a circular-loop, which in turn generates the reflected circular polarized wave. This behaviour is observable since the

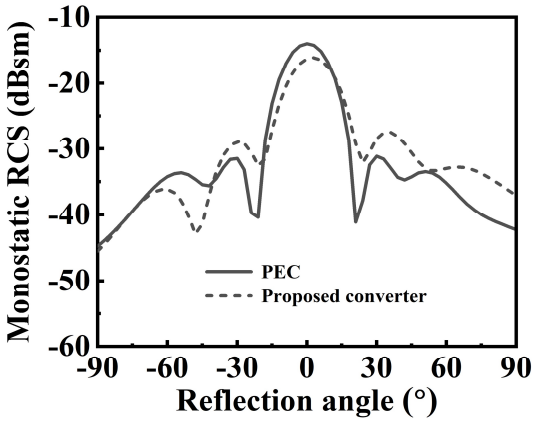


Fig. 10. Monostatic RCS analysis of the proposed design in comparison with PEC.

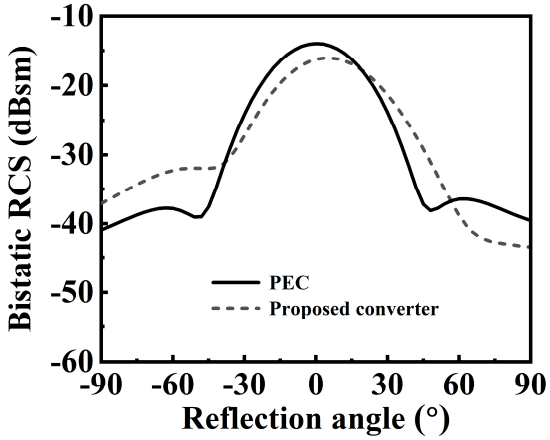


Fig. 11. Bistatic RCS analysis of the proposed design in comparison with PEC.

resonant frequency 15.5 GHz is near to wideband performance of the circular response (see Fig. 3(a)). The current distribution at 7.6 GHz is exhibiting both electric and magnetic response, which means the currents at the lower left of the circular arc are parallel with ground currents. At the lower-right, these currents are anti-parallel with the ground. Multiple plasmonic resonances (both electric as well as magnetic) are responsible for the wideband polarization conversion phenomena.

V. Application of the Proposed Design for Radar Cross Section Reduction

Radar Cross Section (RCS) of an object is the effective area that can be seen by the radar. Since the metallic conductors perfectly reflect all the incident EM energy, RCS is more for the metallic object. Hence, the reduction of Radar Cross Section (RCS) from metallic targets is the main objective in military and defence applications. It can be reduced in many ways, but using planar polarizer reduces the design complexity and make it compact. An analysis has been carried out for proposed design at 10.1 GHz (radar frequency in X-band) to explain the RCS reduction ability of the proposed design. The monostatic and bistatic RCS of 4 x 4 unit-cell array (with dimensions 37.6 mm x 37.6 mm) is compared with PEC, and the simulated results are shown in Fig. 10, and Fig. 11, respectively.

Table. I. State-of-art-performance comparison table.

Reference	Centre frequency (GHz)	Conversion involved	FBW (%)
[12]	20.18	Linear-Cross	77%
[21]	13.55	Linear-circular	30.3%
[22]	9.30	Linear to cross	59.6%
	14.65	Linear to circular	13%
This work	**4.56**	**Linear to cross**	**3.07%**
	8.93		**40.76%**
	15.93		**12.5%**
	4.93	**Linear to circular**	**2.83%**
	6.27		**5.9%**
	12.93		**17.01%**

From Fig. 10, it seems that monostatic RCS of the proposed converter is reduced by 2.3 dB in comparison with PEC at the incident angle θ_i =0° and ϕ_i = 0° (for the vertically polarized EM wave). The non-uniform RCS response due to the asymmetry of the unitcell. As the angle of incidence increases, the RCS of the proposed converter reduces with respect to PEC. Finally, the bistatic RCS study with respect to PEC has been carried out, as shown in Fig. 11. The performance of the proposed converter is compared with the recently reported designs in the literature, listed in Table I. To the best of author's knowledge, the proposed design can fit into the current state-of-the-art table.

VI. Conclusion

In this article, wideband and high-efficiency reflective type polarization converter based on metasurface is demonstrated, which converts the linearly polarized EM wave to its orthogonal, and circularly polarized reflected wave in different frequency bands. A simple dual-cut circular split ring printed on 3.2 mm FR-4 grounded dielectric substrate is used. Fabrication of the proposed metasurface will be done using a traditional printed circuit board (PCB) technology. The compactness of this design, low-profile nature, multi-polarization conversion in different frequency bands makes this device to be used many real-time applications where both linear-cross and linear-circular conversion are needed in various bands.

ACKNOWLEDGEMENT

The authors would like to express their profound gratitude to the Director for setting up Central Instrumentation Facility (CIF) at Indian Institute of Technology Palakkad. This work is supported by Science and Engineering Research Board (SERB), Govt of India under project no: ECR/2018/002258.

References

[1] M. Long, W. Jiang, and S. Gong, "Wideband RCS Reduction Using Polarization Conversion Metasurface and Partially Reflecting Surface," *IEEE Antennas Wirel. Propag. Lett.*, 2017, doi: 10.1109/LAWP.2017.2731862.

[2] N. Yu *et al.*, "A broadband, background-free quarter-wave plate based on plasmonic metasurfaces," *Nano Lett.*, vol. 12, no. 12, pp. 6328–6333, 2012, doi: 10.1021/nl303445u.

[3] J. Ghosh, S. Ghosal, D. Mitra, and S. R. B. Chaudhuri, "Mutual coupling reduction between closely placed microstrip patch antenna using meander line resonator," *Prog. Electromagn. Res. Lett.*, 2016, doi: 10.2528/PIERL16012202.

[4] A. J. Danner, T. Tyc, and U. Leonhardt, "Controlling birefringence in dielectrics," *Nat. Photonics*, 2011, doi: 10.1038/nphoton.2011.53.

[5] J. B. Pendry, "Negative refraction makes a perfect lens," *Phys. Rev. Lett.*, 2000, doi: 10.1103/PhysRevLett.85.3966.

[6] W. Cai, U. K. Chettiar, A. V. Kildishev, and V. M. Shalaev, "Optical cloaking with metamaterials," *Nat. Photonics*, vol. 1, no. 4, pp. 224–227, 2007, doi: 10.1038/nphoton.2007.28.

[7] J. Chen *et al.*, "Observation of the inverse Doppler effect in negative-index materials at optical frequencies," *Nat. Photonics*, 2011, doi: 10.1038/nphoton.2011.17.

[8] J. ming Hao, M. Qiu, and L. Zhou, "Manipulate light polarizations with metamaterials: From microwave to visible," *Front. Phys. China*, 2010, doi: 10.1007/s11467-010-0005-y.

[9] Y. Cheng, Y. Nie, Z. Cheng, and R. Z. Gong, "Dual-Band Circular Polarizer and Linear Polarization Transformer Based on Twisted Split-Ring Structure Asymmetric Chiral Metamaterial," *Prog. Electromagn. Res.*, vol. 145, no. February, pp. 263–272, 2014, doi: 10.2528/pier14020501.

[10] J. K. Gansel *et al.*, "Gold helix photonic metamaterial as broadband circular polarizer," *Science (80-.).*, 2009, doi: 10.1126/science.1177031.

[11] M. Feng *et al.*, "Broadband polarization rotator based on multi-order plasmon resonances and high impedance surfaces Broadband polarization rotator based on multi-order plasmon resonances and high impedance surfaces," vol. 074508, 2013, doi: 10.1063/1.4819017.

[12] X. Gao, X. Han, W. P. Cao, H. O. Li, H. F. Ma, and T. J. Cui, "Ultrawideband and High-Efficiency Linear Polarization Converter Based on Double V-Shaped Metasurface," *IEEE Trans. Antennas Propag.*, vol. 63, no. 8, pp. 3522–3530, 2015, doi: 10.1109/TAP.2015.2434392.

[13] J. Ding *et al.*, "Efficient multiband and broadband cross polarization converters based on slotted L-shaped nanoantennas," *Opt. Express*, vol. 22, no. 23, p. 29143, 2014, doi: 10.1364/oe.22.029143.

[14] F. Li *et al.*, "Compact high-efficiency broadband metamaterial polarizing reflector at microwave frequencies," *IEEE Trans. Microw. Theory Tech.*, vol. 67, no. 2, pp. 606–614, 2018.

[15] Q. Zheng, C. Guo, G. A. E. Vandenbosch, P. Yuan, and J. Ding, "Ultra-broadband and high-efficiency reflective polarization rotator based on fractal metasurface with multiple plasmon resonances," *Opt. Commun.*, 2019, doi: 10.1016/j.optcom.2019.05.044.

[16] X. Jiang, P. Zhong, Q. Zhang, and A. K. Rashid, "A Broadband Metamaterial Polarization Converter Based on Split Ring Resonators," *2019 Cross Strait Quad-Regional Radio Sci. Wirel. Technol. Conf. CSQRWC 2019 - Proc.*, pp. 1–3, 2019, doi: 10.1109/CSQRWC.2019.8799331.

[17] M. S. Ibrahim, A. Mahmoud, A. Awamry, Z. H. Jiang, W. Hong, and M. Al-Nuaimi, "Design and fabrication of engineered reflector for wideband linear-to-circular polarization converter," *2019 IEEE Int. Symp. Antennas Propag. Usn. Radio Sci. Meet. APSURSI 2019 - Proc.*, pp. 1697–1698, 2019, doi: 10.1109/APUSNCURSINRSM.2019.8888310.

[18] F. Costa *et al.*, "Broadband linear to circular reflection polarization converter," *2019 IEEE Int. Symp. Antennas Propag. Usn. Radio Sci. Meet. APSURSI 2019 - Proc.*, vol. 5, pp. 17927–17937, 2019, doi: 10.1109/APUSNCURSINRSM.2019.8888804.

[19] C. Mao, Y. Yang, X. He, J. Zheng, and C. Zhou, "Broadband reflective multi-polarization converter based on single-layer double-L-shaped metasurface," *Appl. Phys. A*, vol. 123, no. 12, p. 767, 2017.

[20] Q. Zheng, C. Guo, and J. Ding, "Wideband metasurface-based reflective polarization converter for linear-to-linear and linear-to-circular polarization conversion," *IEEE Antennas Wirel. Propag. Lett.*, vol. 17, no. 8, pp. 1459–1463, 2018, doi: 10.1109/LAWP.2018.2849352.

[21] Z. Zhang, X. Cao, J. Gao, and S. Li, "Broadband metamaterial reflectors for polarization manipulation based on cross/ring resonators," *Radioengineering*, 2016, doi: 10.13164/re.2016.0436.

[22] Q. Zheng, C. Guo, and J. Ding, "Wideband metasurface-based reflective polarization converter for linear-to-linear and linear-to-circular polarization conversion," *IEEE Antennas Wirel. Propag. Lett.*, vol. 17, no. 8, pp. 1459–1463, 2018, doi: 10.1109/LAWP.2018.2849352.

BitMint Hard Wallet: Digital Payment without Network Communication.

NO INTERNET, YET SUSTAINED PAYMENT REGIMEN BETWEEN RANDOMNESS-VERIFIABLE HARD WALLETS.

Keywords: tamper-resistance, rock-of-randomness, digital money, TEE

Gideon Samid

Department of Electrical Engineering and Computer Science
Case Western Reserve University, Cleveland, OH

Gideon@BitMint.com

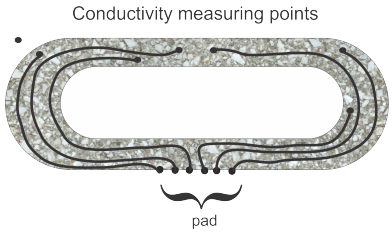
Abstract: Identity-bearing digital money (e.g. BitMint) can be paid in a private transaction between payor and payee without reliance on network authentication. The payee must trust that the paid digital coin is bona fide. Presenting a hard wallet (HW) to generate the required trust. Payment issued from the HW can be taken in by a second hard wallet, which will further pay to a third hard wallet, creating a payment ecology of digital money for long periods without the benefit of a communication network. Payment may be tethered to eventual terms of redemption. The hard wallet may be personalized -- fitted with ownership security capability. The HW may be engineered in conjunction with a smart phone, so people can use a single device as a phone and as an off-line wallet.

I. INTRODUCTION

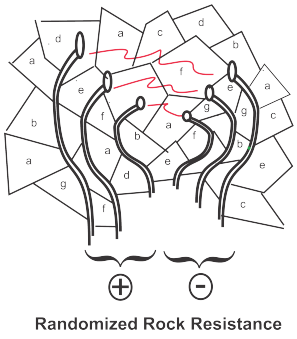
Both legacy money and peer-to-peer digital money are network dependent. A transaction occurs between a payor and a payee, authenticated by a network authentication agent. When the network is absent, or seriously jammed, payment stops. Since payment flow is the blood flow of the body of society, it is important to insure uninterrupted payment regardless of the availability of a communication network serving as a payment authentication agent.

In a tight trust-bearing community such payment can be carried out on the basis of the prevailing trust among members of the community. Trust may be extended via insurance, but a robust payment regimen must allow two strangers to pass money to each other in a trustworthy manner.

The idea raised here to enable such a network-absent transaction is anchored on the notion of the Hybrid Coin [1]. The idea is a physical device that upon inspection by the payee, it generates trust. This physical appearance based trust is the traditional mode of payment for thousands of years before digital money. Payees rubbed banknotes, checked them up towards the sun; coins were bitten to assert their metal constituency, and gold was tested to withstand acids. The hybrid coin, though, is a holder of digital money, denominated as a fixed some inscribed on the surface of a physical device, where by passing it, payment is done. The physical integrity of the coin is the basis of its projected trust. The payee could hold the hybrid coin, or passing it further, or they could crack the shell of the coin, take out the electronic device with the digital bits on it, and incorporate these bits in a digital money environment.



This hybrid coin solution may be upgraded by building a device that is upgraded to a wallet. Namely a physical device that would contain a certain amount of digital money, X, and would pay an amount Y which ranges from some minimum amount to X: $0 < Y \leq X$. Upon examination of the wallet, to be regarded as "hard wallet" (HW), the payee will trust the digital money passed to them from the wallet.



In its basic mode, the HW is one-way. It pays out pre-installed money, but it cannot be reloaded. In the advanced mode the HW will operate two-ways. The idea: if a nominal payee can trust money dispensed from a HW, then a second HW will also so trust. So it is sufficient for a payee HW to confirm that the source of the money bits is another trusted HW, for the payee HW to trust those money bits, and therefore use them as payment, much as the pre-loaded bits are used for payment. In other words, the two-ways payment mode is based not only the singular trust of the one-way paying HW, but on the combined trust of all the participating HW where the money was flowing in that off-line mode.

So in summary, the digital money stored in the HW will be either (i) placed in the HW by the wallet manufacturer, or (ii) passed to the wallet from another qualified hard wallet.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

The community-trust HW payment option leads to a gathering of mutually mistrustful parties, each holding a HW, allowing for any which way payment to support normal societal activity, and all that through periods that may be quite extended when for whatever reason the Internet is not readily available. This off-line payment mode is the missing ingredient on the road for worldwide adoption of digital money in the framework of national fiat currencies.

Since a wallet is an essential take-along device, and so is the phone, then it makes great sense to engineer them jointly. This option is addressed herein.

Ahead we describe (i) architecture, (ii) dynamics, (iii) construction, then discuss tethering, loyalty money application, a Hard Phone Wallet.

II. Hardware Integrity Technology Options

This proposition has two elements: innovative technology to secure physical integrity, and innovative applications to use such technology. The two elements may be regarded in separation. So while the description of the applications regards particular hardware integrity technology, it can be exercised with different technologies.

Here below we survey the technologies considered for hardware integrity. We distinguish between (i) hybrid coin technology, [1] (ii) tamper-resistant technology [2], and (iii) "Rock of randomness" technology [3] Their cost is different, their efficacies are different, and one should match the technology to the application, to the denomination of the coin, and to the risk environment.

III. HW Ecology

HW payment is seen on one hand as a replacement for customary cash payment, and on the other hand as a payment continuity solution for periods (which may be extended) when no network connection is available at the required level.

Much as in the original days of paper money when it was a claim check against gold, so the HW paid digital money is regarded as a claim check against the more common, more historic form of money. Cash payment was traditionally anonymous payment. So it is, or may be, if so designed, for the HW payment. The payee may not care who the payor is, he only cares that the money paid is bona fide, in other words, is redeemable for more legacy cash.

For OW HW (One Way Hard Wallet) the cash payment is terminal, when all the wallets are empty. there is no more activity. But with TW-HW (Two Ways Hard Wallet) the trade can go on forever, as long as the traders have confidence that the transacted digital money is eventually redeemable at par value. It is important to note that HW payment is well fitted into a global digital money ecology, in particular money that bears identity.

The ecology where HW payment is happening depends on the money used. If used for generic fiat national currency, then the range of payment is world wide, in the most general sense, and in that case HW payment is very much like cash.

If used for tethered money then the tethering features may restrict the trading community. Say only members of a club or of a given community can redeem the money, and then the payees will be limited de facto to this community. If the HW paid money is tethered to be redeemed only by grocery stores, then the payees will be restricted to those who wish to use the money for groceries. [5].

Tethering may impose complicated and abstract restrictions. [5] If the HW paid money is only redeemable after a certain future date then, the trade is limited to people who can forgo redeeming the money in the present. If the money is redeemed upon the occurrence of certain future conditions, then the trade is limited to risk taking investors.

IV. One-Way Hard Wallet (OW-HW)

A one-way hard wallet (OW-HW) is a physical device pre-loaded with a sum of digital money, X, equipped with a mechanism to pay off successive amount Y_1, Y_2, ... Y_n where:

$$Y_i \leq X - \Sigma Y_j \text{ for } j=1,2,..(i-1)$$

for i = 1,2,...n.

The payment can be carried out through any means used to transfer digital money. The payor is the HW, but the payee may be any device, or facility fitted to accept bit-wise money.

The payee will be given the opportunity to examine the HW, be satisfied of its bona-fide, and accept the passed bits as due payment. Eventually the paid digital money will be submitted for redemption at the redemption center. If it turns out that the bits do not represent good money then, the payor will be subject to criminal investigation like any money counterfeiter.

The OW-HW cannot be reloaded. Once all its pre-loaded money is exhausted the HW can make no more payments.

A. Architecture (OWHW)

The OWHW system is a configuration of the following items:

1. The HW capsule 2. Authentication Pad (optional) 3. Payee device

The HW capsule is comprised of:

1. HW Shell 2. Internal Electronics (data storage, processor) 3. Payment Ports 4. Human-Machine Interface (HMI) 5. Battery Pocket, standard battery, or recharge port. 6. Ownership Security features (optional)

The HW operation relies on the integrity of the shell, and the trust it generates. Such trust may be based on physical examination, as the case is with the hybrid coin, or it may be

based on a sophisticated test of integrity and validity conducted by the authentication pad, on which the HW wallet is laid (or which is fitted into the HW). The connection between the HW and the payee device may be physical or through any short range electromagnetic communication (e.g. NFC, Bluetooth, WiFi). Alternatively it may be carried out through screen picture (QR or barcode). The HW will erase the paid digital coin from its memory to prevent any possibility of double payment. The money is kept in permanent (e.g. flash) memory so it retains its value even if the battery is removed.

The HW may include any of the tamper resistance measures mentioned in ref [2] or others. However, by erasing the money bits that are paid out, any physical compromise of the HW will not enable double payment.

The shell is fitted with a human machine facilitation (Human Machine Interface) where the owner of the wallet may both read the digital content of the wallet, as well as any relevant information, and at the same time, the owner can tell the wallet to pay any amount available in its memory, and otherwise communicate with the HW.

The wallet, optionally can be fitted with ownership security features designed to insure that the wallet will only be used as a payment source by its rightful owner.

B. Payment Dynamics (OWHW)

We consider a payor holding a OW-HW and a payee holding a payment accepting device, geared for digital money. The two parties agreed on a sum Y to be paid from the payor to the payee. They decide to exercise the payment off-line.

The payor will pull out the HW and present it to the payee. If the sums are sufficiently small, then the payee, upon cursory physical examination of the wallet will agree to consider payment from it, as bona fide. The payor will show on the HMI screen that the wallet contains money in the amount $X \geq Y$, and so it may serve as a bona fide source of payment. In other modes the payor does not have to expose the value of Y. The payee will check to see if they got the expected amount Y, once it has been transferred to them.

If the HW is fitted with anti-theft ownership security features, then the rightful owner will activate these features, to activate the HW. If the HW holder fails to satisfy the ownership security features, then the HW does not dispense money.

Next the payee will use the HMI interface to set up the payment action. The payee device will be set up to receive the money and OK to the payor that the payee device is set, then the payor will click the pay button on the HMI, and the wallet will dispense the indicated sums through the chosen communication channel to the payee device.

The payee will find out from their device that the money has been paid, report this to the payor, and the payor will so indicate on the HMI, and in response the electronics inside the HW will erase the bits corresponding to the paid amount. Alternatively the money bits can be erased as soon as they are communicated outside the HW.

In the event that the sums involved are large or otherwise the payee is not comfortable accepting the risk of physical examination of the wallet as described, then the payment will require a technology based authenticity test. Upon its satisfactory execution the transaction will be carried out.

In the case where the wallet is fitted with ownership validation means, then they would have to be activated and allow the wallet owner to operate.

This nominal payment dynamics may be extended to automatic payment.

It is very important for the HW to erase paid money bits (payload). This will de-motivate anyone from cracking the HW and double spending the money in it. In most protocols the paying HW will wait for the recipient device to OK the reception of the money, but such may be skipped, for certain versions of the protocol. A recipient device may fail to OK receipt of money, or the transferred money bits will be taken in by a device that is not a wallet. In that case the device may register the sent over bits but will be unable to process them and verify that those bits constitute a valid digital coin according to the protocol of the mint of that coin. In BitMint, for example, the digital coin has a well-defined structure and a distinction between meta data and payload data, and every wallet software will recognize a BitMint coin for what it is. However, a bit recipient which is not a wallet will not be able to recognize the coin for what it is, and hence will not be able to OK the payment to the payor HW. The payor HW will nonetheless erase the paid bits to absolutely prevent double spending.

C. Automatic Payment

The OW HW may be fitted to dispense money automatically. A payment management software will dictate successive amounts Y_1, Y_2, ... to be paid per controlling events. Such events may be a simple passage of a time interval, like in paying for parking or for time measured services, or they may be specified events that the payee is undergoing or the environment is undergoing. The idea here is that the payments are carried out fast over an established channel. In some application payment may be effected through just erasing of money bits without passing them out to anyone.

In a nominal mode payor and payee will cooperate in a dialogue as discussed above, for any passage of money. In the automatic mode the dialogue will be established once, and then the money will flow as the events dictate.

D. Payment Channel

The actual payment dynamics will depend on the channel chosen to effect the coin delivery. We discuss: (i) hardware connection, (ii) Bluetooth, (iii) NFC, (iv) Camera-QR.

Hardware connection may be accomplished to standard USB and micro USB ports, or through specialized ports. It may be accomplished through direct touching or through connectivity pads on payor and the payee device, or it can be done via a connecting cord. The passage of the digital coin may follow any common relevant protocol. Same for the Bluetooth, the NFC and the camera readout. The industry developed standard communication protocols for each mode.

E. Construction (OWHW)

We discuss construction of the OW HW elements:

1. The HW capsule 2. Authentication Pad (optional) 3. Payee device

The hard wallet will accommodate all common means for passing bits, so all money handling computing devices can serve payees in this context.

F. The HW Capsule

The HW capsule is comprised of:

1. HW Shell 2. Internal Electronics (data storage, processor) 3. Payment Ports 4. Human-Machine Interface (HMI) 5. Battery Pocket, standard battery. 6.Ownership features (optional) .

G. The Hard Wallet Shell

The Hard Wallet Shell is the trust generating element in the hard wallet. Simple versions, use the Hybrid Coin technology [1], extending its use from a coin of fixed denomination to a wallet from which money can be paid in any amount up to the contents of the wallet. The idea being that any attempt to crack the shell will be noticed by the payee. In other words, the assumption is that it would be too difficult, too costly for a fraudster to either crack the wallet, or otherwise tamper with the content. The next level is the technology identified in [2], where any optional combination of light, pressure and internal composition of a gas, will erase the data upon tampering, and no payment can be made.

The higher up mode is based on the rock of randomness technology identified in [3]. This technology is further innovated to establish high level of security and integrity (and hence trust) for the wallet.

H. Rock of Randomness Shell Technology

One constructs the shell according to the technology identified in [3] namely as a randomness guided 3D printed combination of conductivity-variant elements. At one flat location on the shell (from the outside) a plate is fitted with t ports. Each port is the end of a highly conductive well-insulated wire that is fitted inside the shell layer and ends at a random location inside that shell, where the wire is exposed (no insulation on the end point of the wire). The wires may either be fitted as part of the 3D printing of the 'rock' or threaded in after the rock is 3D printed. The wire connects between a port on the surface of the shell, and a randomized location in the shell, where the wire terminates.

So fitted one could randomly pick p < t ports and attach them to one pole of a battery, then combine other ports, one, or up to all the remaining q= t-p ports and attach the picked ports to the other pole of a battery, and thereby measure the effective resistance between these specified two groups of wire terminal points spread throughout the shell of the HW. The shell is identified with a unique ID affixed to it by the manufacturer. An authentication agent can compare the resistance reading from the shell to a reading of the same as registered in a ledger prepared by the manufacturer of the shell. If the two readings data agree, then there is a high likelihood that the measured shell is the shell that was manufactured and documented as to per its various conductivity readings from various combinations of ports. To increase the likelihood for a bona fide, non counterfeit shell, one would pick another combination of p' ports and a respective combination of q' where q' is comprised of one port up to q' = t- p' ports, and compare again the actual reading of the shell to the recorded matching reading issued by the manufacturer. The manufacturer publishes the ledger while the Internet is on and working. The HW and any other computing devices, like the authenticators, all download the master file, the ledger published by the manufacturer, and this copy of the ledger they use it to confirm HW (shell) identity in periods without a functioning Internet.

To exercise this test one uses an authentication device (authentication pad). It also stores a copy of the manufacturer database (ledger)which features assorted conductivity readings of each of the shells manufactured by the manufacturer. Owing to this pre-stored database, the pad can compare the reading done over a given shell to the readings stored in tis memory and if they check out, to issue a statement of authentication.

I. HW Internal Electronics

Inside the shell there is (i) a permanent bit storage device, e.g. flash memory, and (ii) a processor. The memory holds the money bits, and auxiliary data. The processor has read/write access to the memory element, and it has access to the ports of payments through which the money bits flow to the recipient. The processor also is communicating with the HMI interface to (a) report to the wallet operator, and (b) accept payment instructions and other instructions from them. If Ownership security elements are installed, the processor connects to them too, and will only operate if the security check if positive.

V. HUMAN-MACHINE INTERFACE

The Human Machine interface (HMI) element will be following the common practice of smart phones, where the screen is used both as mode of data entry and mode of data display. Used for the wallet operator to give payment instructions or other instructions, and used for the wallet to report the amount of money available at any moment in the wallet, as well inform the operator on any related aspects. In particular, if the money in the wallet is tethered and the

request for payment does not satisfy the term of payments. The wallet will keep track of its activity. How much money was paid when, and if known, to whom. Using the BitMint money protocol, the HW will retain the capsule of the coin, namely the coin metadata (while the payload, the actual money bits are erased once paid).

VI. Authentication Pad (optional)

The authentication pad applies for shells based on the 'rock of randomness' technology. These shells will have a tablet displaying an array of contact points (conductivity measuring ports), where each port is the end of an insulated conductivity wire, which terminates somewhere inside the 'rock material' of the shell. The tablets with the n contact points is fitted over the authentication pad which is fitted with a matching array of electrical terminal points.

The tablet on the shell also includes an ID terminal, identifying electronically the identity of the particular shell. The tablet on the pad has a matching reader terminal to read the ID of the shell.

There are there types of authentication pads:

1. online authentication pad 2. off-line authentication pad 3. on-and-off line authentication pad

The online pad has an online link to the HW manufacturer database. Based on the pad reading the shell ID, it queries that manufacturer database to receive a conductivity reading based on some chosen combination of p conductivity port (out of the array of n) to be assigned to one electrical pole, and the combination of the remaining q (where q comprises 1 port up to q = n – p ports) conductivity points (ports), grouped into the opposite pole. The authentication pad will then measure the conductivity of the shell with the same pole (electrode) grouping. If the result is a match then the shell is declared authentic. The pad will have an authenticity indicator to light up to communicate this conclusion. This indicator can be in a form of a light -- green for Ok, red for not-OK: fraud suspected.

The off-line pad will have to be pre-loaded with the manufacturer database (ledger) where the full set of wallets that have a chance to be engaged at that location, is identified, each with its own ID and a series of conductivity measurements based on some particular grouping of the n conductivity ports. Reading the ID of the wallet submitted for authentication, the authentication pad will query the downloaded database for conductivity data for that ID, and again compare the stored results to actual measurement of the wallet submitted for authentication. The report likewise, will be submitted via the authentication indicator. If the ID of the submitted wallet is not found in the downloaded database, the pad will indicate: "cannot authenticate the submitted wallet".

The on and off line pad, will operate online if online connectivity is available, and will resort to off line mode otherwise.

VII. OW-HW Security

OW-HW security may be viewed from (i) the payee stand point, and (ii) from the HW owner stand point. The payee is only interested in the question of whether the money paid to it is bona fide. This assessment will be made by the payee based mainly on the sum paid and on the hardware security grade. Small daily amounts are a good risk to be taken from low level HW since the effort to forge such a wallet is greater than the transacted value. For higher sums, a higher degree of HW integrity will be required. The rock technology is considered sufficiently secure to guarantee the bona fide status of even millions of dollars HW paid.

The security considerations of the HW owner are different. Since HW payment needs to be paid while exposing the wallet, the owner stands the risk of being robbed. The security answer to this risk is combined ownership verification features.

We distinguish between (i) biometric security, (ii) PIN security, and (iii) ledger security.

Biometric security is any biological signature of the owner. E.g. finger prints, facial recognition, iris, palm, etc. There can be a few of them in either "OR" or in "AND" mode.

The PIN security is based on entering a secret PIN. The PIN can be a fixed string of bits, or it can be dynamic. For example it can be time dependent: every day it changes, and the owner needs to memorize the new PIN every day, based on applying a secret app on his phone or computer.

Ledger security is based on the HW ledger identifying not only the conductivity readings (in the Rock technology) but also the owner of the HW. The payee will then ask the HW holder to identify himself per the indication on the ledger.

Any combination of these features is acceptable.

The selected security features will be built into the shell of the HW.

The HW is programmed to erase the money bits after paying them, so that any one who cracks the shell will not be able to make a double payment from money that was already paid. In the BitMint format the HW may keep the coin 'capsule' for any paid digital coin. The capsule is the full meta data of the coin, in fact, all the data of the coin except the payload. This way the HW keeps a full record of its payments over time, but keeps no copy of money that was paid once, denying a shell breaker the opportunity to profit from double payment, and hence deny him the incentive to crack the HW shell in the first place.

Biometric data of the owner may be entered once, when the owner first receives the hard wallet. That data is then kept inside the HW and cannot be changed, to enhance security.

VIII. Two Ways Hard Wallet (TW-HW)

In the one-way HW mode, the wallet cannot be recharged. Its content is drained until it is empty. It would be much

more viable to exploit the deep trust engendered by the hard wallet to practice a two-ways transaction mode.

In the one-way mode the recipient, without a HW will trust the money bits passed to them from a one-way HW on the basis of trusting the physical hard wallet. But these bits once held outside a hard wallet, cannot be simply paid to a stranger under conditions of no-network. BitMint developed means for off-line payments that involved cryptographic trust mechanism, but they are involved, and require payor and payee to come from the same trading environment where these means are accepted and recognized. They also operate for a limited period of time. When it comes to a general transaction between complete strangers, the payee is worried about (i) the authenticity of the passed digital coin and about (ii) whether some double spending is involved. This apprehension is calmed down when the payor is a HW, but when the payee of a OW-HW tries to use the money given to it, for its own payment, then the new (2nd) payee re-emerges with the above mentioned apprehension because the paid money is not secured in a HW. This means that in a trading community, once the money stored in a set of OW-HW is paid out, then no more ready payment under no-communication regime can take place. Say then that a OW-HW runs its course when all wallets are empty, and if the network is not back by then -- trade stops.

This terminal status of OW-HW is a fundamental issue. The two-way mode is designed to resolve it. A two-way HW is built much as a OW-HW is built only that it allows for re-loading. This re-loading will be allowed only if the incoming money will enjoy the same security projected on outflowing money. This equivalence of security can be achieved by limiting the source of the incoming money to be an equally secure HW wallet. In other words, if the rock technology is good enough to convince a stray payee that the money that is coming out of a HW is bona fide then it is also good enough for a HW to accept such money coming directly from a similar HW unit.

All that is needed is for the payee HW (the reloaded unit) to make sure that the source of the inflowing money bits is no other than a similarly secure HW. To achieve this confidence by the payee HW, it would be necessary for the payee HW to ascertain the HW status of the payor. This can be done either (i) directly, or (ii) indirectly -- using an authentication pad.

A. Direct HW2HW Authentication

The two involved HW will be physically connected via an array of s electronic ports. $n < s$ of these ports will be the rock conductivity ports, and the remaining $m = s-n$ will be used to transfer money and data.

In one design out of many, the connection will be facilitated via a connection pad fitted on the back of each TW-HW. The two HW will be placed back to back. The following will happen:

1. The payor HW will send the payee HW (i) the ID of the payor HW, and the amount of money it wishes to pay. This information can also be passed via Bluetooth or other means.

2. The payee HW will consult its built-in, previously loaded (or real time loaded, if the network is operational) ledger of ID data, and pull out of it, conductivity data with respect to the identified payor ID. The payee HW will pull first one recorded resistance, (conductivity) measurement, R(p,q), where R is the measurement per a particular selection of p conductivity ports and q conductivity ports, where $p+q \leq n$ the full range of conductivity ports. The TW HW will have the built in circuitry to conduct a live measurement of the conductivity reading per the identified p and q ports off the payor HW. If the readings don't match -- the dialogue is terminated. If the readings match then the payee HW is developing confidence that the device attached to it, ready to make a payment, is indeed the HW it claims to be. The payee HW then signals to the payor HW: 'exercise payment!'.

3. The payor HW will then send the money bits and any related information per the particular protocols. These bits and information will flow through the m ports on the pad.

4. For the payor's HW, this payment is no different than payment to any payee device. But for the payee HW, this incoming money is now added to the sums previously stored in the payee HW, only that a record is kept of the ID source of that money (along with timing, and perhaps location information – of GPS data is available).

The just added money is now imbued with full trust because the payor HW was duly authenticated. That means these money bits can be paid out with the same reliance and confidence as the pre-stored moneys. Optionally the data regarding the chain of holding HW devices may be kept.

Note that for this direct mode to work the HW will have to house the HW resistance readings ledger. Also note that the authentication protocol may call for two or more different resistance readings with different combinations of p and q ports.

B. TW-HW Security

Security for TW-HW follows the logic and the features, of the OW option. In addition the TW version has a verification and authentication pad to allow a 2nd HW to verify and authenticate a payor hard wallet, (1st HW), and also allow a third HW to verify and authenticate the 2nd hard wallet, and so on. If an intermediate security pad is used it will also be rock-of-randomness protected. So each hard wallet will trust the other HW through trusting the authentication pad.

IX. HARD TETHERED WALLET (HTW)

Both the one-way and the two-ways modes for hard wallet operate in the basic mode on untethered money. Namely the owner of the money in the wallet can pay to anyone,

anytime, for any reason, like any money owner. This technology can also be used to handle tethered money, namely money that can be paid only subject to pre-set restrictions.

The payment restrictions (the money tethering) will be inscribed in the wallet to insure that any payment is carried out in compliance of the pre set restrictions. (The money will not be eventually redeemed if the tethering terms have not been complied with). Nominally tethered digital money is controlled through its redemption process, where the redeemer will only redeem a submitted digital coin, if it satisfies the tethering terms. However the hard wallet is designed to work in conditions where the network and hence the redeemer are not accessible, and therefore the enforcement of the tethering is left to the hard wallet. That means that the payor HW will dispense of the money only if the payee HW has supplied proof or indication that it is wallet eligible to receive the tethered money. This will require an additional dialogue between the payor and the payee HW.

Loyalty Money: Merchants could choose the HW technology to dispense their store loyalty money. The HW can be given anonymously, and it can be given to customers to auto-pay. The HW may be latched to a payment pad on a 'wall' where the money will be quickly paid without any need for online authentication which may slow things down. This "pay into the wall" idea can be used for after hours purchases from a side wall exposed to the public. Payment will allow a cell door to open and for the merchandise to be picked up.

Crowd Movement Payments: The HW may be used by a flow of people going into a train station or alike. They will attach the HW to a payment pad on a wall and pay cash like in great speed.

X. HARD PHONE WALLET (HPW)

The hard wallet may be combined with a smart phone to construct a unit that is both a smart phone and a hard wallet. This will spare users the need to hold and take along two items. The Hard Phone Wallet (HPW) will function as both a phone and a wallet. The rich phone interface will serve the wallet for all transactions. The integrity of the physical unit will generate the necessary trust to spur back and forth payments.

An HPW may be constructed using the Hybrid coin shell technology, for small amounts of money. But it can be constructed with 'rock' technology to turn the phone into full fledged secure banks. The HPW may be built in (i) the inclusive mode and in (ii) the latched mode. In the former case the secure shell will engulf everything except the display and communication screen. In the latter case the wallet will be the shell-secured unit, latched permanently or not, to the phone back to back. The wallet and the phone will be communicating with each other through a communication pad fitted with communication ports. That way the screen of the phone can be used to communicate with the wallet.

The HPW will have its own authentication pad built in, as described in the TW-HW, so that one HPW can securely pay to another HPW, and a full ecology of payment can be practiced, all the while without the benefit of a communication network.

XI. COMPARATIVE DISCUSSION

The issue known as TEE (Trusted Execution Environment) is addressed by major manufacturers like Intel's SGX [5]. Smart phones use SIM cards and ARM to create "Trust Zones" [6]. A notch above is provided by Hardware Secure Modules (HSM) which claim tamper resistance at a rather high cost. Lately secure hardware for bitcoins were proposed [7] . Ideas to prevent double spending have attracted attention. E.g. Zcash [8]. Zcash is a complex implementation and not very versatile. It is generally agreed that a well motivated, well financed attacker can compromise the various ideas presented in the literature. So most solutions are practically limited to small sums. By contrast the BitMint Hard wallet presented here features hardware that is manufactured through quantum-grade randomness, and no known technology can reverse engineer the wallet from a given set of its measurements. And since the user can increase the number of measurements that characterize the hard wallet shell, the user is in control of how much security the hard wallet projects. For hard wallets containing large sums of money the security will be properly upgraded. Since the software in the hard wallet erases any paid bits, then there is no incentive for an attacker to break the hard wallet shell. In summary it appears that the proposed BitMint hard wallet is a viable solution for the challenge of sustained off line payment of digital coins.

XII. REFERENCE

1. US PATENT 9,471,906
2. US PATENT 10,445,730
3. US PATENTS, 10,467,522
4. SAMID, WNEK, "THE "ROCK OF RANDOMNESS": A PHYSICAL ORACLE FOR SECURING DATA OFF THE DIGITAL GRID" HTTPS://WWW.CAMBRIDGE.ORG/CORE/JOURNALS/MRS-COMMUNICATIONS/ARTICLE/ROCK-OF-RANDOMNESS-A-PHYSICAL-ORACLE-FOR-SECURING-DATA-OFF-THE-DIGITAL-GRID
5. Intel Software Guard Extensions, Reference Number: 332680-002, 2015. [Online]. Available: https://software.intel.com/sites/default/files/ 332680-002.pdf.
6. T. Alves and D. Felton, TrustZone: Integrated Hardware and Software Security-Enabling Trusted Computing in Embedded Systems, 2004. [Online]. Available: http://infocenter.arm.com/help/topic/com.arm.doc.prd29-genc- 009492c/PRD29-GENC-009492C_trustzone_security_whitepaper.pdf.
7. Bitcoin Wiki, Hardware wallet, 2020. [Online]. Available: https://en.Bitcoin.it/wiki/Hardware_wallet.
8. E. Ben-Sasson, A. Chiesa, C. Garman, M. Green, I. Miers, E. Tromer, and M. Virza, "Zerocash: Decentralized anonymous payments from Bitcoin", in Security and Privacy (SP), 2014 IEEE Symposium on, IEEE, 2014, pp. 459– 474.

Design of Pyramidal Horn Antenna at 7.5 GHz by using different Optimization Techniques

***Soham Kanti Bishnu, Sutapa Ray, Malay Gangopadhyay**

Department of Electronics and Communication Engineering (ECE)
Institute of Engineering and Management.
Kolkata, India.
*soham.kanti.bishnu@iemcal.com

***Abstract*- In this work, we discuss about the design of a pyramidal horn antenna with high gain and reduced side lobes partially covering the C band and the X band of operation. The waveguide width (a), waveguide height (b), flare width (A), flare height (B) and flare length of the pyramidal antenna have been optimized. These dimensions determine the required characteristics such as S11, gain, radiation pattern of the antenna. The optimized antenna shows decent gain of approximately 12.2 dB over operating range while having a bandwidth of 3 GHz centered at the frequency of 7.5 GHz (approximately 7 GHz to 10 GHz). Ansys HFSS has been used to design and simulate the antenna. Genetic Algorithm, Quasi-Newton method and Pattern search optimization techniques are comparatively utilized to optimize the antenna design.**

Keywords- Pyramidal horn antenna, HFSS, Genetic Algorithm, Quasi-Newton method, Pattern Search method

1. INTRODUCTION

Pyramidal horn antenna has huge application in the field of communication systems at UHF and microwave frequencies. The advantages of such antenna are that they have high gain and efficiency, the side lobes are suppressed to a very low level and that they can be used at very high frequencies[1]. In this paper, we have demonstrated a design of pyramidal horn antenna at 7.5 GHz frequency in HFSS simulation software and optimized it using three optimization techniques. Optimization techniques used are Genetic Algorithm, Quasi-Newton method, Pattern search techniques. We have analyzed a comparative study of gain and return loss for these three techniques. Using these optimization techniques, we basically optimized five parameters of a horn antenna. They are such as waveguide width (a), waveguide height (b), flare width (A), flare height (B) and horn length.

2. RESEARCH METHODOLOGY

A. Genetic Algorithm

Genetic algorithm is a special type of optimization technique where evolution of solution is observedsuch as competition, fitness based selection, recombination such as cross over and mutation in each generation and if required, more generations is added to get the best solution [2]. Each solution has a set of properties (its gain or S_{11}) which can be mutated and altered.The evolution usually starts from a number of randomly generated parameters, which is an iterative process and these parameters in each iteration is called a generation[3]. In each generation, the solutions are compared among each other and the best solution is selected. Commonly, the algorithm stopsif a particular number of generations have been created or a satisfactory solutionis achieved for the population.

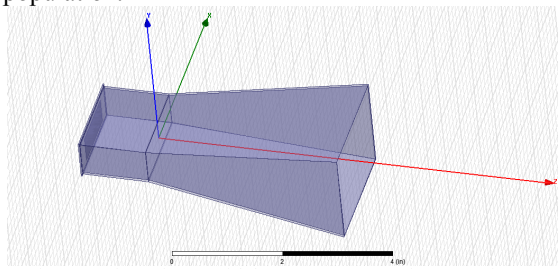


Fig. 1 Design of Pyramidal Horn Antenna using Ansoft HFSS

In present case, Genetic algorithm was given a population size of 30 and allowed to evolve for 100 generations. The calls to the fitness evaluator constitute the majority of the runtime of the algorithm. Approximately the same runtime for both algorithms is indicated with equal number of calls to the fitness evaluator. Various parameters of optimization process are shown in the Table below.

978-1-7281-9615-2/20/$31.00©2020 IEEE

Table 1. Different parameters in Genetic Algorithm

Maximum number of generations:	100
Elapsed time	1 hour
Number of Individuals in Parent Population	30 (Roulette selection)
Number of Individuals in Children Population	30
Parento Front Number of Survivors	10
Number of Individuals in Next Generation	30 (Roulette selection)

B. Quasi Newton Method

Quasi-Newton methods are used to find the maximum and minimum values in a function. Newton's method requires the Jacobian to find minimum value and the Hessian for finding maximum value. If Jacobian or Hessian can't be used, the best alternative method is Quasi-Newton method [4].

Quasi-Newton methods are a special kind of variable-metric methods based on Newton's method to find the minima of a nonlinear function. To search the minimum value Newton's method uses the first and second derivatives and for maximum value it uses the gradient and the Hessian matrix of second derivatives of the function. This method is used to reduce the computation cost and complexity of the Newton method.

C. Pattern Search Method

Pattern search method is a type of numerical optimization methods.It is used on functions that are not continuous or differentiable and does not require any gradient to optimize [5]. This method is also used to optimize the design of the antenna and to get a desirable gain and S_{11}. The parameter values are shown in Table 2.

3. RESULTS AND OBSERVATIONS

The initial measurements and measurements of different dimensions of horn antenna after optimization have been given below in the Table 2.

Table 2.Measurement of the parameters of horn antenna using three different optimization techniques

Name	Original Dimension (in)	Genetic Algorithm (in)	Quasi Newton (in)	Pattern Search (in)
a	1.13	1.63530487990	1.63530314960	1.69180314960
b	0.5	0.60189367351	0.60189370078	0.70189370078
Flare A	2.25	2.18878368480	2.18878346456	2.18878346456
Flare B	1.75	1.75616855372	1.75616929133	1.75616929133
Horn length	3.75	3.95502716147	3.95503937007	3.76753937007

A comparative reading of Gain and S_{11} in different optimization techniques at 7.5 GHz frequencies are shown in below Table 3.

Table 3. Determination of gain and S_{11} using three different optimization techniques.

Algorithm	Gain(dB)	S_{11}(dB)
Genetic Algorithm	12.11176	-28.0296
Quasi Newton method	12.21178	-28.4244
Pattern Search	12.23156	-30.0908

The radiation pattern has been generated in Genetic algorithm is shown in Fig. 1.

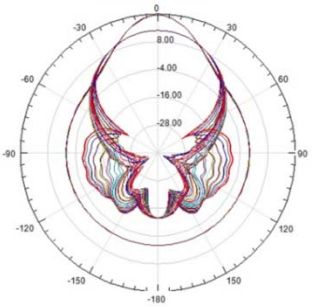


Fig. 2 Radiation pattern after optimizing with Genetic algorithm

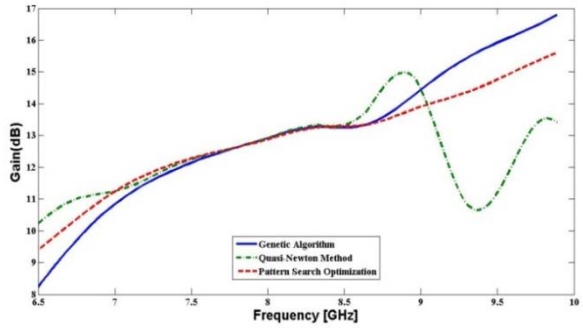


Fig. 3 Comparative study of gain generated in three different optimizing techniques.

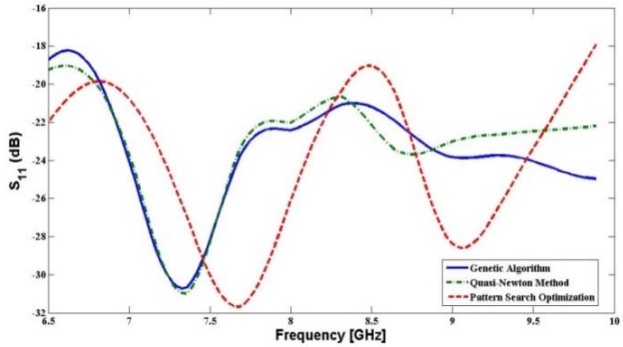


Fig. 4 Comparative study of S_{11} generated in three different optimizing techniques.

4. CONCLUSION

A comparative study of S11 and gain using different optimization techniques has been reported in this paper. The main purpose of this paper is to design a Pyramidal Horn Antenna at 7.5 GHz using HFSS software. We have simulated three different optimization techniques such as Genetic Algorithm, Quasi-Newton method and Pattern search to optimize the antenna design. We have optimized Waveguide width (a), waveguide height (b), flare width (A), flare height (B) and horn length of the pyramidal antenna using these three optimization technique. In Pattern search optimization technique gain is almost flat about the centre frequency while in case of genetic algorithm and Quasi Newton method it steadily increases and varies in sinusoidal manner. On the other hand, we find that the geometry optimized with Pattern search has minimum return loss near the centre frequency as compared with other algorithms. It can thus be concluded that Pattern search is the most preferred optimization technique among these three for designing an optimal horn antenna. The final design shows good performance parameters in terms of gain and bandwidth.

ACKNOWLEDGMENT

We would like to take this opportunity to extend our sincere gratitude to Dr. Kaushik Patra for his continuous, motivation and support.

REFERENCES

[1] Shubhendu Sharma, "Design and Analysis of Pyramidal Horn Antenna at 8 GHz Frequency",International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 3, Issue 2, February 2014.

[2] Jacob Robinson, Seelig Sinton, and YahyaRahmat-Samii,"Particle Swarm, Genetic Algorithm, and their Hybrids: Optimization of a Profiled Corrugated Horn Antenna", IEEE Antennas and Propagation Society International Symposium, DOI:10.1109/APS.2002.1016311

[3] D.S. Weile; E. Michielssen,"Genetic Algorithm Optimization Applied to Electromagnetics: A Review", IEEE Transactions on Antennas and Propagation, Volume: 45, Issue: 3 , Mar 1997, Pages 343 – 353.

[4] Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, Numerical Recipes, Cambridge University Press, New York, 2003.

[5] FilizGünes, FikretTokan,"Pattern Search optimization with applications on synthesis of linear antenna arrays", Expert Systems with Applications, Volume 37, Issue 6, June 2010, Pages 4698-4705.

[6] J. Robinson, S. Sinton and Y. Rahmat-Samii, "Particle swarm, genetic algorithm, and their hybrids: optimization of a profiled corrugated horn antenna," IEEE Antennas and Propagation Society International Symposium (IEEE Cat. No.02CH37313), San Antonio, TX, USA, 2002, pp. 314-317 vol.1.

[7] H. Faezi and J. R. Mohassel, "Genetic Algorithm Optimization of Corrugated Horn Antennas," 2005 Pakistan Section Multitopic Conference, Karachi, 2005, pp. 1-4, doi: 10.1109/INMIC.2005.334460.

[8] Nicholas V. Findler, Cher Lo, Ron Lo, "Pattern search for optimization", Mathematics and Computers in Simulation, Volume 29, Issue 1, 1987, Pages 41-50, ISSN 0378-4754, https://doi.org/10.1016/0378-4754(87)90065-6.

[9] R. Krishna Chaitanya, G.S.N. Raju, K.V.S.N. Raju & P. Mallikarjuna Rao, "Antenna Pattern Synthesis Using the Quasi Newton Method, Firefly and Particle Swarm Optimization Techniques", IETE Journal of Research, DOI: 10.1080/03772063.2019.1643263

[10] T. Hassan, A. Kausar, H. Umair and M. A. Anis, "Gain optimization of a seven element ESPAR Antenna using Quasi-Newton method," 2011 IEEE International Conference on Microwave Technology & Computational Electromagnetics, Beijing, 2011, pp. 293-296, doi: 10.1109/ICMTCE.2011.5915515.

[11] M. Abbas-Azimi, F. Arazm, J. Rashed-Mohassel & R. Faraji-Dana, "Design and Optimization of a New 1–18 GHz Double Ridged Guide Horn Antenna", Journal of Electromagnetic Waves and Applications, 21:4, 501-516, DOI: 10.1163/156939307780616810

Optimal Design of Antenna Array using Tuned Random Drift Particle Swarm Optimization Algorithm

Debolina Brahma
School of Electronics Engineering
Kalinga Institute of Industrial Technology, Deemed to be University
Bhubaneswar, India
debolina.brahma90@gmail.com

Arindam Deb
School of Electronics Engineering
Kalinga Institute of Industrial Technology, Deemed to be University
Bhubaneswar, India
arindamdeb2004@yahoo.co.in

***Abstract*—Conventional particle swarm optimization and random drift particle swarm optimization algorithms are applied to the optimal design of a collinear dipole antenna array. The random drift particle swam optimization algorithm is tuned for its best performance by properly choosing the control parameter and it showed a faster convergence rate compared to the conventional particle swarm optimization algorithm. The optimal array designed using the conventional particle swarm optimization algorithm has a beam width of 9 degrees between 1st nulls and maximum side lobe level of -24.17 dB whereas the array designed using the random drift particle swarm optimization algorithm has a beam width of 10 degrees between 1st nulls and maximum side lobe level of -27.46 dB.**

Keywords—random drift, particle swarm optimization.

I. INTRODUCTION

Particle Swarm Optimization (PSO) and its improved versions are a popular choice among antenna engineers for optimizing their designs to suit multiple objectives for various applications. Extensive application of the PSO algorithm to electromagnetics was initially reported in [1]. Since then, various improved PSO algorithms have been proposed for better performance in antenna and array design problems [2]-[4]. Improved and different versions of evolutionary optimization algorithms have been applied to different antenna array synthesis in the recent past [5]-[8]. Random Drift Particle Swarm Optimization (RDPSO) was analyzed in [9] and applied to power economic dispatch problem in [10]. A modified version of the RDPSO algorithm was successfully applied to target tracking problem in [11]. It appeared to be a very promising algorithm in the optimization of antenna arrays in [12] where flat top pattern synthesis using RDPSO was reported by using a single run of the algorithm without any parameter tuning. In this paper, the suitability of the RDPSO algorithm with respect to the design of a collinear dipole antenna array design is reported and its statistical performance is compared with that of the conventional PSO algorithm. The RDPSO algorithm is tuned for optimal performance by suitable choice of control parameter.

This paper is organized as follows. Section I provides the introduction to the paper. Section II gives an overview of the conventional PSO and RDPSO algorithms. Section III explains the optimization problem, associated cost function and numerical set-up for performing statistical study. Section IV explains the tuning of the RDPSO algorithm and compares the statistical performance of the two algorithms. Section V concludes the paper.

II. CONVENTIONAL AND RANDOM DRIFT PARTICLE SWARM OPTIMIZATION ALGORITHMS

A. Conventional Particle Swarm optimization Algorithm

The conventional Particle Swarm Optimization (PSO) algorithm operates by assuming a collection of bees or swarm in the solution space whose positions indicate possible solutions to the optimization variables defined for a particular optimization problem [1]. The coordinates of the bee's position indicate a particular cost value. The bees move in different directions based on velocity and position update equations controlled by their own local best positions and the global best position of the entire swarm. The velocity update equation along each dimension for each bee is given below:

$$v_{n+1}(i,j) = w.v_n(i,j) + C_1.rand().\left(P_n(i,j) - x_n(i,j)\right) + C_2.rand().\left(G_n(i,j) - x_n(i,j)\right) \quad (1)$$

where $v_n(i,j)$, $x_n(i,j)$,$P_n(i,j)$ *and* $G_n(i,j)$ represent the current velocity, current position, local best position and global best position along the j^{th} dimension for the i^{th} bee in the n^{th} generation. C_1, C_2 *and* w represent the local acceleration coefficient, global acceleration coefficient and inertial weight for the swarm. The local best positions and the overall global best position play a crucial role since they control the movement of the bees in every iteration.

The position update equation along the j^{th} dimension for the i^{th} bee in the n^{th} generation is given by:

$$x_{n+1}(i,j) = x_n(i,j) + v_{n+1}(i,j) \quad (2)$$

B. Random Drift Particle Swarm Optimization Algorithm

The Random Drift Particle Swarm Optimization (RDPSO) algorithm operates by assuming flow of a collection of electrons in a conductor under the effect of an externally applied electric field [10]. The positions of the electrons represent possible solutions to the optimization problem. Each electron is assumed to have two velocity components i.e. the drift velocity and the thermal velocity.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

The velocity update equation for the drift velocity component is given below:

$$v_{D(n+1)}(i,j) = \beta(P_n(i,j) - x_n(i,j)) \quad (3)$$

where β is the drift coefficient and $P_n(i,j)$ is the local focus of each particle as given in [10].

The velocity update equation for the thermal velocity component is given below:

$$v_{T(n+1)}(i,j) = \alpha|C_n(i,j) - x_n(i,j)|\lambda_n(i,j) \quad (4)$$

where α is the thermal coefficient, $C_n(i,j)$ is the mean best position [10] and $\lambda_n(i,j)$ is a random number having standard normal distribution. The position update equation is given by (2).

III. Cost Function Formulation and Numerical Set-up

A. Optimization Problem and Cost Function Formulation

Optimal pattern synthesis for a linear array of collinear dipole antennas is investigated in this paper. The optimal pattern is expected to have narrow Beam Width between 1st Nulls (BWFN) and minimum Side Lobe Levels (SLL). The collinear dipole array arrangement is shown in the Fig. 1 below. 16 antenna elements are considered and they are assumed to be symmetrically excited about the center of the array.

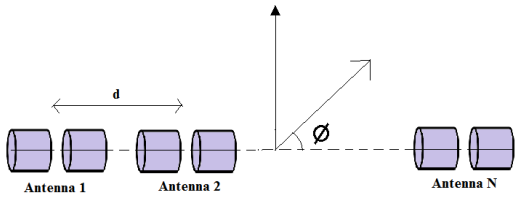


Fig. 1. Linear array of collinear dipole antennas

The pattern of a single dipole element is given by the following equation:

$$E = j\frac{\eta I_o e^{-j\beta r}\cos\left(\frac{\pi}{2}cos\varphi\right)}{2\pi r\, sin\varphi} \quad (5)$$

where η, Io, r and φ represent the impedance of free space, the excitation current amplitude, far field distance and scan angle respectively. The array factor of the equivalent linear array with isotropic elements is given by:

$$AF = |A_o + A_1 z^1 + A_2 z^2 + \cdots \ldots\ldots\ldots\ldots z^{n-1}| \quad (6)$$

where,

$$z = e^{j\psi}, A_m = I_m e^{j\alpha_m}\ , \psi = \beta d cos\varphi$$

The overall pattern is obtained through pattern multiplication and the normalized electric field amplitude is designated as $E_n(\varphi)$. The dipole elements of the array are to be excited in such a way so that the optimization objectives are achieved within a predefined number of iterations of the optimization algorithms considered in this paper. Besides, the inter-element spacing between the collinear dipole elements is also optimized to obtain good results.

The cost function for this optimization problem is formulated as:

$$Cost = C_1\frac{1}{E_n(\varphi=90^o)} + C_2 BWFN(radians) + C_3\max(SLL) \quad (7)$$

where C_1, C_2 and C_3 are unity. The first term ensures that the principal maxima occurs at φ=90 degrees. The second and third terms in the cost function ensure that BWFN and maximum side lobe level are minimized to the maximum extent possible.

B. Numerical Set-up

Performance comparisons of conventional particle swarm optimization and random drift particle swarm optimization algorithms are done with respect to the optimization problem defined in the previous sub-section. Each of the optimization algorithms is run for 20 independent trials for 150 iterations starting with the same initial population. The population size is taken to be 40 for both the algorithms. There are 17 optimization variables; the optimization variables include the current amplitude and phase excitation of the first 8 elements of the array; the inter-element spacing, 'd' is optimized within the range 0.5λ-1λ.

The local and global acceleration coefficients for the conventional particle swarm optimization are taken to be 2.0. For the random drift particle swarm optimization, the parameter that needs to be tuned is the drift coefficient, β. The thermal coefficient, α is varied linearly from 0.9 to 0.3 over the number of iterations [10]. In this paper, the performance statistics of the RDPSO algorithm are computed for all values of β ranging from 1.1 to 2.0 at intervals of 0.1 to identify the value which gives the best statistical performance. The performance data is given the next section.

IV. Results and Observations

The statistical performances of the conventional PSO and RDPSO algorithms with different values of β are given in the table below:

TABLE I. Performance Comparison of Optimization Algorithms

Optimization algorithm	Mean cost value	Best cost value	Standard deviation of the cost values
PSO	1.2243	1.2189	0.0047
RDPSO(β=1.1)	1.2518	1.2257	0.0239
RDPSO(β=1.2)	1.2491	1.2206	0.0237
RDPSO(β=1.3)	1.2348	1.2199	0.0126
RDPSO(β=1.4)	1.2378	1.2168	0.0238
RDPSO(β=1.5)	1.2399	1.2157	0.0212
RDPSO(β=1.6)	1.2286	1.2171	0.0172
RDPSO(β=1.7)	1.2259	1.2169	0.0067
RDPSO(β=1.8)	1.2297	1.2193	0.0107
RDPSO(β=1.9)	1.2331	1.2194	0.0088
RDPSO(β=2.0)	1.2445	1.2210	0.0337

The performance of the RDPSO algorithm is seen to be best with β=1.7.The average convergence plots for the

conventional PSO and the RDPSO (with β=1.7) algorithms are shown below in Fig. 2:

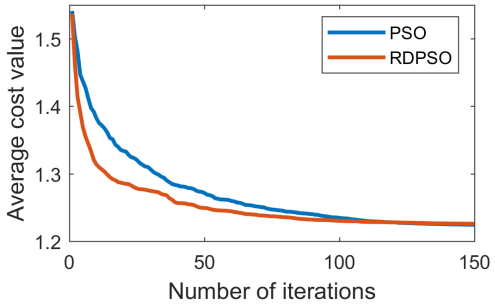


Fig. 2. Variation of average cost value with iterations.

The statistical performance of the tuned RDPSO algorithm with β=1.7 is almost same as that of the conventional PSO algorithm. However, as seen in Fig 2, the convergence rate of the tuned RDPSO algorithm is much faster compared to the PSO.

The excitations of the optimal arrays from the best run of the PSO and tuned RDPSO algorithm are given in the figures below:

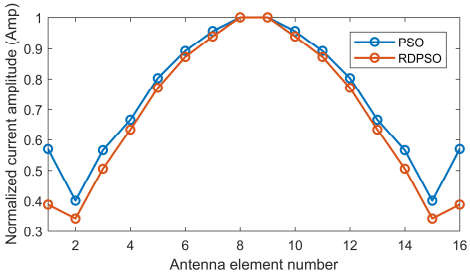


Fig. 3. Normalized current amplitude excitations of the optimal arrays from the best run of the PSO and tuned RDPSO algorithm.

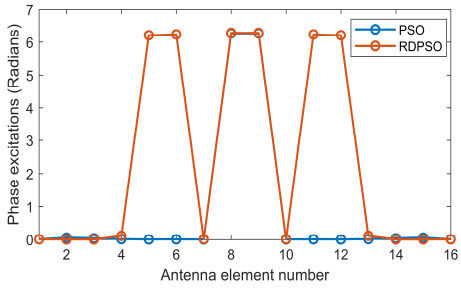


Fig. 4. Current phase excitations of the optimal arrays from the best run of the PSO and tuned RDPSO algorithm.

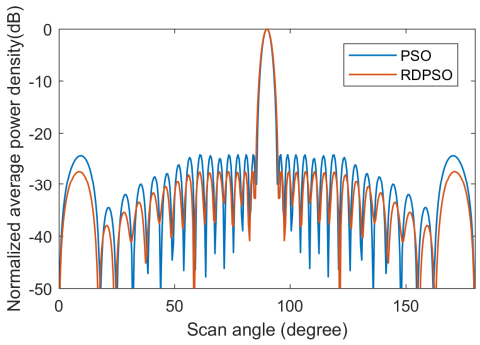


Fig. 5. Normalized power pattern of the optimal arrays.

The normalized power patterns of the optimal arrays corresponding to the excitations shown in Fig. 3 and Fig. 4 are depicted in Fig. 5 above.

The characteristics of the optimal arrays are given below:

TABLE II. PERFORMANCE COMPARISON OF OPTIMAL ARRAYS

Optimization algorithm	BWFN(degrees)	Maximum SLL (dB)	Inter-element spacing
PSO	9.00	-24.17	0.97λ
RDPSO(β=1.7)	10.00	-27.46	0.96λ

V. CONCLUSION

Performance of the random drift particle swarm optimization (tuned for optimal performance) is statistically compared with the conventional PSO algorithm. It is observed that the RDPSO algorithm performs best with β=1.7. Further, the characteristics of the optimal arrays obtained from the best run of the two algorithms are also compared. The inter-element spacing in case of both the arrays is almost same. In case of RDPSO optimized array, the beam width between the 1st nulls (i.e. 10 degrees) is slightly more compared to the PSO optimized array (BWFN=9 degrees), but the maximum side lobe level in case of the former (-27.46 dB) is significantly lower compared to the latter (-24.17 dB).

REFERENCES

[1] J. Robinson and Y. Rahmat-Samii, "Particle swarm optimization in electromagnetics," *IEEE Transactions on Antennas and Propagation,* vol. 52, no. 2, pp. 397-407, 2004.

[2] N. Jin and Y. Rahmat-Samii, " Advances in particle swarm optimization for antenna designs: Real- Number, binary, single-objective and multiobjective implementations," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 3, pp. 556-567, 2007.

[3] H. Wu, J. Geng, R. Jin, J. Qiu, W. Liu, J. Chen and S.Liu, "An improved comprehensive learning particle swarm optimization and its application to the semiautomatic design of antennas," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 10, pp. 3018-3028, 2009.

[4] R. Bhattacharya, T.K. Bhattacharyya and R. Garg, "Position mutated hierarchical particle swarm optimization and its application in synthesis of unequally spaced antenna arrays," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 07, pp. 3174-3181, 2009.

[5] F. Liu, Y. Liu, F. Han, Y. Ban and Y. Jay Guo, "Synthesis of large unequally spaced planar arrays utilizing differential evolution with new encoding mechanism and Cauchy mutation,"*IEEE Transactions on Antennas and Propagation*, vol. 68, no. 6, pp. 4406-4416, 2020.

[6] R. Wang and Y. Jiao, "Synthesis of sparse linear arrays with reduced excitation control numbers using a hybrid cuckoo search algorithm with convex programming,"*IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 3, pp. 428-432, 2020.

[7] Z. Qi, Y. Lu, Y. Kong and Y. Cong, "Hybrid optimisation algorithm and its application for pattern synthesis of planar arrays,"*The Journal of Engineering*, vol. 2019, no. 20, pp. 6612-6615, 10 2019.

[8] C. Zhang, X. Fu, X. Chen, S. Peng and X. Min, "Synthesis of uniformly excited sparse rectangular planar array for sidelobe suppression using multi-objective optimisation algorithm,"*The Journal of Engineering*, vol. 2019, no. 19, pp. 6278-6281, 10 2019.

[9] J. Sun, X. Wu, V. Palade, W. Fang and Y. Shi, "Random drift particle swarm optimization algorithm: convergence analysis and parameter selection," *Machine Learning*, vol. 101, pp. 345-376, 2015.

[10] J. Sun, V. Palade, X-Jun Wu, W. Fang and Z.Wang, "Solving the power economic dispatch problem with generator constraints by random drift particle swarm optimization," *IEEE Transactions on Industrial Electronics*, vol.10, no. 1 ,pp. 222-232, 2014.

[11] W. Ding and W. Fang, "Target tracking by sequential random draft particle swarm optimization algorithm,"*2018 IEEE International Smart Cities Conference (ISC2)*, Kansas City, MO, USA, 2018, pp. 1-7.

[12] B. Misra and A. Deb, "Synthesis of antenna arrays with flat-top pattern using conventional and random drift particle swarm optimization algorithms,"*2018 Emerging Trends in Electronic Devices and Computational Techniques (EDCT)*, Kolkata, India, 2018, pp. 1-5.

Wi-Fi Based Indoor Positioning and Navigation System (IPS/INS)

Jaren Golenbiewski
Electrical and Computer Engineering
Kettering University
Flint, USA
jarengski@gmail.com

Girma Tewolde
Electrical and Computer Engineering
Kettering University
Flint, USA
gtewolde@kettering.edu

***Abstract*— Over the last several years there has been fast growing research activities in the areas of indoor positioning, tracking, navigation, and their applications. Commercial, industrial, and retail businesses are highly motivated in developing solutions for accurate localization of assets and people. This paper first presents a brief review of the various technologies for indoor positioning applications, and then chooses one of them, the Wi-Fi RSSI based technique, for further development of a prototype application at a university campus. The algorithms that are commonly used for calculating the user's position are also presented, including Trilateration, Fingerprinting, and the K-Nearest-Neighbor (KNN). For the prototype implementation, an indoor positioning and navigation system is developed with an Android app that any user can utilize to navigate through the target indoor environment. At the end, the paper presents evaluation results of the prototype implementation, which demonstrates the effectiveness of this developed low-cost solution for indoor positioning and navigation applications.**

Keywords*— *Fingerprinting, Indoor Navigation System (INS), Indoor Positioning System (IPS),* K-*Nearest-Neighbor (KNN), Trilateration, Wi-Fi

I. INTRODUCTION

With the proliferation of Smartphones over the last dozen of years, GPS based positioning, navigation, and tracking services have become ubiquitous. The satellite based GPS system works well for the most part in outdoor environments, however, within indoor spaces, GPS signals may be unavailable or become too weak to be usable. It is therefore impossible to rely on GPS for navigation in indoor environments, where people spend most of their time to work and live. Shopping centers, airports, schools, university campuses, hospitals, and museums are just some examples where indoor positioning would be able to bring much needed benefit to people.

Users would like to have access, on their mobile devices, to indoor maps marked with their current location and navigational services that facilitate seamless movement to specific locations of their interest. This technology definitely revolutionizes indoor navigations. Such services could also bring additional benefits to organizations and businesses, from delivery of location-aware services to the offering of location-triggered content and location-based targeted advertising.

Similar to satellite based systems that provide position references for enabling GPS receiver devices to calculate locations, most indoor positioning technologies today rely on some form of beacon sources to serve as anchor points for the localization algorithms. The beacon sources would typically rely on one or a combination of the following technologies for providing reference signals: Bluetooth Low Energy (BLE), Wi-Fi, ultrasonic, infra-red (IR), Ultra-Wideband (UWB), Visual Light Communication (VLC), and Geomagnetic signals.

The rest of this paper is organized as follow. Section II presents the technologies that are used in most of the current implementations of the indoor positioning and navigation applications. The prototype development and implementation of a Wi-Fi RSSI (Received Signal Strength Indicator) based system for indoor positioning in our department is presented in Section III. Section IV presents the algorithm developed for path planning to guide users navigate from their initial position to a destination point. Analysis of the project results are presented in Section V. Finally, Section VI presents conclusions of the paper.

II. INDOOR POSITIONING TECHNOLOGIES

This section presents a brief review of the technologies which are commonly used for indoor positioning systems (IPS) and indoor navigation systems (INS). The complete study with analysis of the different technologies is available in the authors' previous paper [1].

A. Bluetooth Low Energy (BLE)

A BLE beacon is a small low-power device that uses radio signals to broadcast advertising packets at customizable intervals to other devices in the surrounding. These packets can be received by a smart mobile device and analyzed to calculate the approximate distance between the transmitting and receiving devices. If packets from three or more BLE beacons are received by the user's device, the distance information can be used by a trilateration technique to calculate the user's location. This technique requires that the exact placement of the BLE beacons is known, which is leveraged in the position calculation.

Two major categories of beacons, namely iBeacon [2] and Eddystone [3] are currently popular in the market. These BLE

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

beacons utilize software SDKs to extract the radio signal strength indicator (RSSI) measurements for the calculation of distance estimates [4].

B. Wi-Fi

Wi-Fi access points (APs) can be utilized as beacon sources in a similar way as BLE beacons for indoor localization. There are a few approaches for utilizing Wi-Fi received signal strength indicator (RSSI) values to estimate the user location. Since each technique has its own benefits and drawbacks, it could be difficult to identify the best technique for a given indoor environment. Typically, choices are made based on the environment configuration as well as the application use case.

1) Trilateration

A trilateration technique using Wi-Fi RSSI measurements is similar to that of the BLE beacon based implementation, except that instead of BLE beacons, Wi-Fi APs are used as beacon signal sources. After receiving RSSI values from three nearby APs, the user's smart mobile device can convert the RSSI values into distance estimates using an RF signal propagation model. It can then apply the trilateration technique which is based on simple distance calculations to determine the user's location. A detailed explanation of the trilateration algorithm is available in [5].

The trilateration technique assumes that the distance measurements are accurate, however, in reality the RSSI to distance conversion, which follows a log-distance path loss model [5] is prone to errors, because of the non-linear relationship between the two and a number of factors that influence the signal strength of the Wi-Fi signal, such as multi-path propagation and interference by other radio signal sources sharing the same frequency spectrum in the environment. Therefore, the trilateration technique may not offer the best position accuracy because obtaining accurate distances from the RSSI measurements is more complicated than it seems.

2) Fingerprinting

In the fingerprinting technique, the incoming RSSI values of the Wi-Fi access points are utilized for positioning by comparing the measurements at the current location with stored RSSI data collected at known reference points. It utilizes statistical relationships to come up with a good estimate of the user's position.

This fingerprinting based positionining technique consists of two phases: training and positioning. The objective of the training phase is to build a fingerprint database for the operating environment. To generate a database, a select set of reference points (RPs) must first be carefully identified to provide sufficient coverage of the operating environment. A mobile user then moves to those selected RP locations one at a time and collects the RSSI values from each AP. This measurement is stored in the database to provide the fingerprint dataset.

During runtime, in the positioning phase, as the user navigates through the environment the user's mobile device collects the RSSI values from its nearby APs. These measurements are then compared with the previously collected fingerprint data in the database using an appropriate search or statistical matching algorithm. The generated result then becomes the location estimate for the user [6].

Note that this method does not require the development of an accurate RF signal propagation model. It avoids this by simply using the RSSI values directly, unlike the trilateration technique which converts the RSSI values to distances. The drawback of the fingerprinting based positioning technique is that it takes a significant amount of time to perform the offline data collection stage. Moreover, in dynamically changing indoor environments the movement of large furniture or machinery could negatively impact the positioning accuracy. To minimize such impacts the fingerprinting data may need to be updated as the indoor environment changes which could add to the cost of the system.

3) *k*-Nearest-Neighbor (KNN)

The *k*-Nearest-Neighbor (*KNN*) algorithm is a widely used technique for classification and other related applications. The main step in this algorithm involves the selection of the nearest *k* neighbors around the mobile user to determine its position [7]. For a Wi-Fi RSSI based positioning implementation, the *k*-nearest-neighbors are selected based on the RSSI values. After these values are sorted, the APs with the highest RSSI values are chosen as reference points for the position calculation. The value of *k* is typically selected experimentally by running multiple tests in the indoor environment. Other approaches, such as the one proposed by [7], modify the basic technique to enhance the *KNN* algorithm by dynamically adjusting the appropriate value of *k*.

Some implementations of *KNN* use weights to enhance the influence of close neighbors, these algorithms are called Weighted *K*-Nearest-Neighbor (*WKNN*). Because of its improved performance over the basic *KNN*, the *WKNN* technique is the one selected for prototyping the indoor positioning system at our university campus. The implementation details are provided in Section III.

C. Other Technologies

Even though BLE beacons and Wi-Fi are currently the most commonly used technologies for indoor positioning systems, researchers continue to explore alternative solutions. Such other options include geomagnetic [8], visual light communication (VLC) [9], ultra wideband (UWB) [10], and wearable sensors [11]. A more detailed examination and relative comparison of these techniques is provided in [1].

From the relative comparison of the different technologies, currently the Wi-Fi based solution offers one of the lowest deployment costs [1]. This is mainly because it can be developed using existing Wi-Fi infrastructure within a building to support wireless network access, thus there are no additional costs incurred to support the IPS. Similarly, unlike the other technologies, scaling the Wi-Fi based solution would not involve the deployment of additional infrastructure, if there are already existing access points in the environment. Moreover, since the APs are typically connected to the building's electrical power supply, there are no maintenance costs associated with replacement of batteries. One drawback of the Wi-Fi based

solution is its limited accuracy, which on average is currently in the range of several meters.

A BLE beacon system with a high density of beacons could potentially offer a better accuracy than a Wi-Fi based solution. However, it would involve high deployment cost for the purchase of enough Bluetooth beacons to cover the entire indoor environment. At every point where determination of location information is desired, three or more beacons must be within the broadcasting range of the point. Due to their limited broadcasting range, BLE beacons are required every 7 to 10 meters, which means higher density of BLE beacons are required compared to more sparsely spaced Wi-Fi access points. Moreover, since most BLE beacons are battery powered, regular maintenance is needed which would add to the cost and complexity of the system. Thus, BLE based IPS is not as easily scalable as that based on existing Wi-Fi infrastructure.

III. WI-FI RSSI BASED INDOOR POSITIONING SYSTEM

As discussed in the previous section the Wi-Fi RSSI based technology was selected for the development and implementation of the indoor positioning system (IPS) at our University [12]. For the actual implementation, a weighted *k*-nearest-neighbor (*WKNN*) algorithm is developed using an Android app to generate the location estimate of the mobile user in real-time within a predesigned map.

A. WKNN Algorithm

The *WKNN* algorithm implemented in this prototype was inspired by Shin's enhanced weighted *k*-nearest-neighbor algorithm as presented in [7]. The pseudo-code for the algorithm implementation is given below:

1. Read RSSI values from the APs
2. Find the *k* APs with highest RSSI values
3. Use the smallest of the *k* RSSI values found in step 2 as the minimum threshold
4. Calculate the weights for each of the remaining (*k*-1) RSSI values using the following formula:

   $W_n = RSSI_n - Minimum\_threshold$

5. Multiply each AP's (X, Y) location by their calculated weight:

   $Weighted\_X_n = W_n * X_n$

   $Weighted\_Y_n = W_n * Y_n$

6. Calculate the total weights using Equation (4)
7. Calculate the *(x, y)* position estimate of the user by dividing the sum of the weighted coordinate values by the total of the weights, using Equation (5).

To demonstrate the operation of the *WKNN* algorithm, let's consider an example given in Figure 1. There are 5 access points (shown in blue) in the environment. The RSSI values that the mobile user (shown in red) reads from the 5 APs are given. Table 1 displays the (x_n, y_n) coordinates of each AP and the calculated values from the algorithm, for the total weight and the estimated *(x, y)* location of the user.

$$W_t = \sum_{n=1}^{k-1} W_n \tag{4}$$

$$(x, y) = \frac{\sum_{n=1}^{k-1}[W_n * (x_n, y_n)]}{W_t} \tag{5}$$

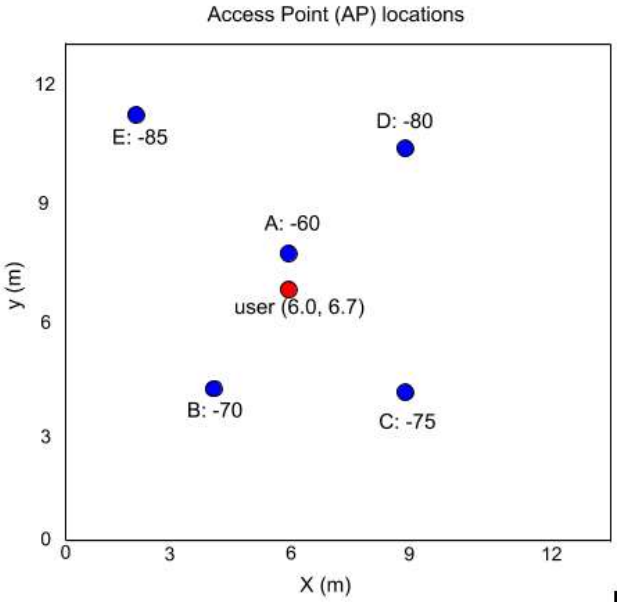


Figure 1. WKNN Example.

Table 1. *WKNN* Algorithm Example Data

AP	X (m)	Y (m)	Weight	Weighted X	Weighted Y
A	6.0	7.8	25	150	195
B	4.2	4.2	15	63	63
C	9.0	4.2	10	90	42
D	9.0	10.2	5	45	51
		Totals	55	438	351
			Final	6.33	6.38

In this example, the true position of the user is (6.0, 6.7). For the sake of simplicity of the example the value of *k* is set to 5, with the five APs shown within the surroundings of the user's location. The user's mobile device performs RSSI scan from all the APs within its close proximity to obtain the values shown in Figure 1. Note that the strength of the RSSI depends on distance and the transmission power of the radio signal source. Higher negative values correspond to relatively lower actual received

power levels. Access point E has the lowest RSSI value at -85, thus this becomes the minimum threshold value, and for the rest of the calculations this AP won't be used any further.

Next, the weights for the access points A through D are calculated as described here. For example, for access point A, evaluate $RSSI_A - RSSI_E = (-60) - (-85)$, which results in a weight value of 25. This process is repeated for all the APs. The next step in the *WKNN* algorithm is to take each weight and calculate the weighted X and weighted Y values. For example, for access point A, the weighed X value becomes 6.0 * 25 or 150. The next step is add all the weighted X's together and all the weighted Y's together, which are then divided by the sum of the weights to obtain the estimated location of the user.

The calculated position of the user generated by the *WKNN* algorithm is (6.33, 6.38), while the actual location of the user was (6.0, 6.7). The result shows a discrepancy from the true location of the user. The main contributing factor for not achieving high accuracy is because the *WKNN* algorithm does not take into account the very complex nature of the propagation of radio signals in indoor environments cluttered with objects and other electromagnetic waves that could potentially interfere with the Wi-Fi signals. The main reason for using this less accurate method is primarily to reduce the implementation cost and complexity.

In the experiments conducted it was learnt that the value of *k* in the *WKNN* algorithm affects the quality of the result. To determine an optimal value for *k*, a test experiment was conducted to calculate a stationary user's position for varying values of *k*. For this exercise, the actual position of the user was first set at a specific known location. The *WKNN* algorithm was then run for varying values of *k* with the location calculated for each case. It was then observed from the results of the experiment that the *k* value of 7 produces the most accurate position for the test compared to other values of *k*. Therefore, for the rest of the project implementation *k* was set to 7 for the *WKNN* algorithm based IPS prototype development.

The value of *k* in the *WKNN* implementation used for this project was determined ahead of time and then held fixed. In alternative approaches, the *k* value could be determined at run time. For example, in the enhanced weighted *k*-nearest-neighbor algorithm [7], the value of *k* is dynamically determined based on a threshold value for eliminating low RSSI signals from consideration for the position estimates.

B. Application Components

The mobile App for the indoor positioning system is developed using Android Studio. The screenshot of the Android app is shown in Figure 2. This section briefly explains the main components of the app.

1) Location indicator

A location indicator on the app is created to show the user's current location, which gets updated dynamically as the user moves around. Due to the possible errors in the location estimate as explained in the previous section, the user's location is represented in the map with a single blue dot surrounded by a

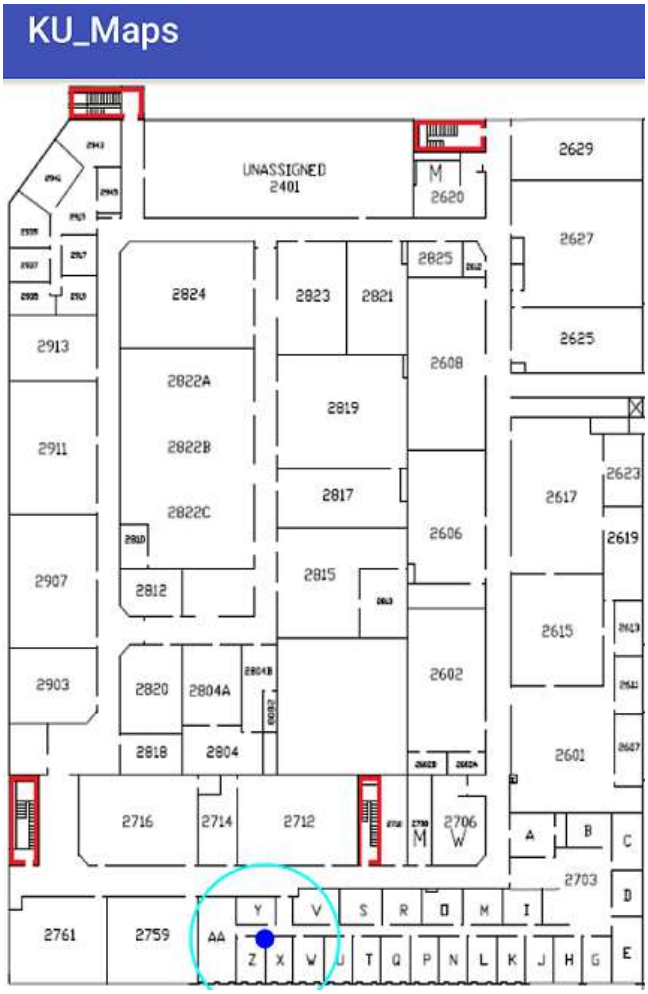


Figure 2. The KU_Maps app, the user is located within the cyan circle.

circle. The size of the circle corresponds to the uncertainty in the position estimate. The mobile application uses information about the display screen size of the smart device to maintain an appropriate and scaled size across different devices and pixel resolutions.

2) Zoom/Scroll

For intuitive user experience, the mobile app offers capabilities for zooming and scrolling features. The zoom scale factor and the transformation matrix of the view are available from the Android APIs. This information coupled with the scaling matrix are used for implementing a transformation function for translating a position in absolute (screen) coordinates to the overall map coordinates.

IV. PATH PLANNING

To support navigational services for mobile users it is important to implement a path planning algorithm. Planning a path for navigation can often be modeled as a search problem in a graph domain, for which a number of graph-based search algorithms have been created for obtaining the least-cost path in the graph with some defined cost parameters. In general, the most common path planning techniques can be categorized into

deterministic, heuristic-based, and randomized algorithms [13]. A-star (A*) algorithm, which is one of the most popular graph-based search techniques for motion planning, was chosen for implementation in this project. A* is founded on the classic Dijkstra's graph search algorithm, with some heuristic- based optimizations applied to improve the performance of the search.

The result generated by the A* path planning algorithm provides an optimal route for a user navigating from their current location to a destination point. The development of the A* algorithm requires breaking up the two-dimensional environment into a map made up of small 2-D grids. The 2-D grids are then represented as graphs, with each block represented as a node in the graph and the connections between neighboring blocks become the edges. If a block contains a wall that block is not navigable and it is excluded from the graph search.

The A* algorithm generally provides the lowest cost path provided that the heuristics for estimating cost measures is defined properly. To simplify the A* implementation, a cost value of 1 was used for each horizontal and vertical movements, and diagonal movements were excluded. The A* algorithm terminates once the destination is found. What makes A* more efficient compared to other similar path planning algorithms is that it does not need to explore and evaluate every possible solution before returning the optimal solution, under the specified heuristics.

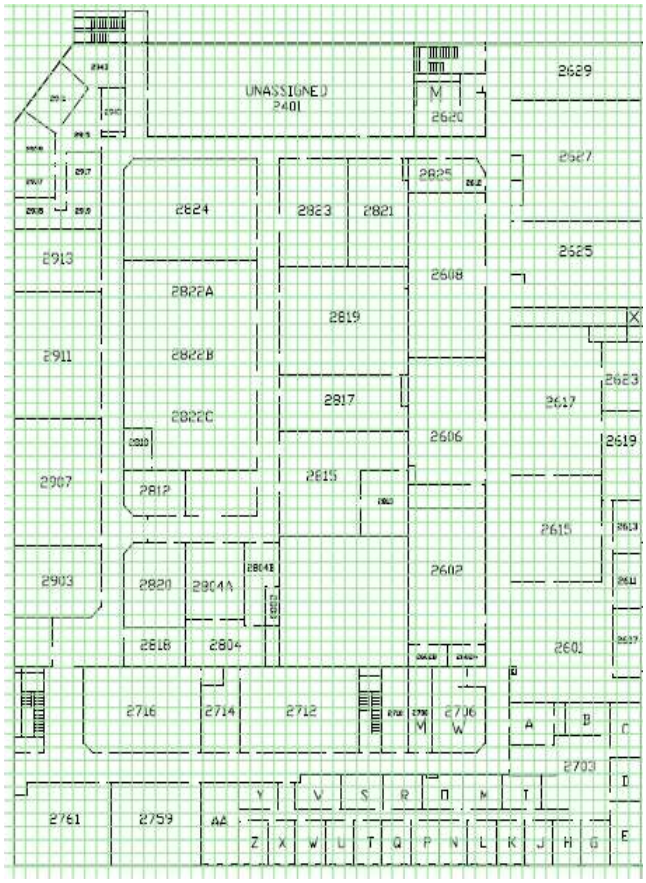


Figure 3. Grid of ECE Department Map

Figure 3 shows the grid representation for the environment under consideration which uses 3,564 blocks, 54 along X and 66 along Y. The map dimension is 65.8 m along X and 80.5 m along Y. Each block is 1.22 m by 1.22 m. This grid provides a resolution that fits one to two blocks in the width of a hallway, which allows for reasonable drawing of the generated path.

A higher resolution map could be used but it would be at the cost of increasing the computational time for the A* algorithm. For this prototype, the chosen resolution performs adequately. The computation is quick enough to not be noticeable by the end user while the resolution is fine enough to provide a path to any designated room on the map. The specifics of the A* algorithm will not be further explained in this paper. However, references [13] and [14] provide good explanations that help understand the basic theory and implementation ideas for the algorithm.

The A* algorithm implemented in this work has been observed to correctly return the least-cost path. Figure 4 demonstrates an illustration of the least-cost path generated by the algorithm from the user's currents position designated by the blue dot inside the cyan circle to the destination room (room # 2818).

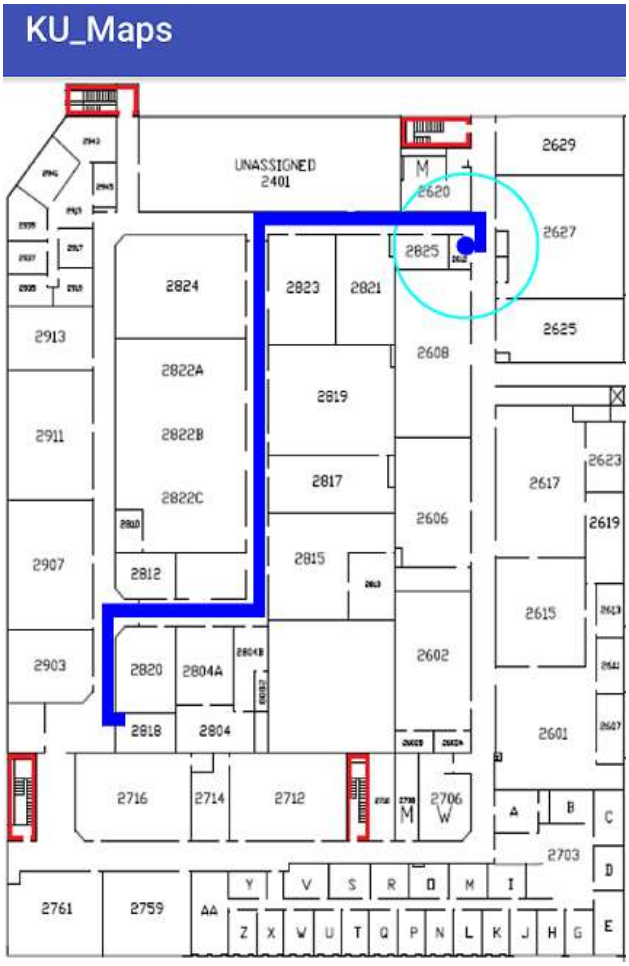


Figure 4. Path plan generated by the A* algorithm.

V. ANALYSIS AND DISCUSSION

The original goals of this indoor navigation project were to meet the following requirements:

1. Create mobile app with intuitive user interface
2. Map representation with zooming & scrolling
3. Show user position with acceptable accuracy
4. Path planning & routing capabilities
4. Scalable
6. Low cost

After the indoor navigation system prototype has been implemented, several experiments were conducted to evaluate its performance. The first three goals could be easily verified to have been successfully met by operating the app with proper user interactions. The app's GUI are as shown in Figure 2 and Figure 4.

To evaluate the localization accuracy, the user walks around to different locations in the environment while observing the calculated positions. These observed positions are then compared against the actual position of the user on the ground. After several experiments at different locations in the environment the average error of the position calculation is found to fall into two categories:

a) Average localization error in the range of 0 to 5 meters. These relatively good localization accuracy correspond to areas with good Wi-Fi coverage by multiple access points, and with little or no obstructions such as walls or other large equipment in the path of the radio signals. Parts of the hallways and areas near some of the classrooms have localization accuracies within this range.

b) Average localization error in the range of 5 to 10 meters. Such relatively lower accuracy correspond to areas that do not receive strong Wi-Fi signals from multiple access points. Also some areas face significant signal degradation from the concrete walls in the building and the heavy machinery inside some of the laboratories.

For improving the localization accuracy it would be possible to apply a different technique for Wi-Fi based localization, such as the fingerprinting approach. This could potentially help keep the error to less than 5 meters. However, the improved accuracy of this alternative method comes at a significantly higher cost in terms of major initial setup task which is needed to collect the fingerprinting data of the environment. Moreover, the fingerprinting data needs to be updated if the environment changes over time.

Another alternative option is to use a combination of multiple technologies that can help improve the accuracy and reliability of the system. For example, to provide additional reference information for the positioning algorithm, Bluetooth Low Energy (BLE) beacons could be used in the areas that have relatively higher inaccuracies due to Wi-Fi weak-spots. However, this method would incur additional cost to deploy the BLE beacons.

For further improvement of the position accuracy the mobile device's built in sensors could be utilized for dead-reckoning. Real-time data from sensors such as accelerometers and gyroscopes can be combined using sensor fusion algorithms to track the device's orientation and movement speed. If the direction and number of steps a user is taking from a reference point is calculated, this information could be used to improve the estimate for a user's location.

The prototype system presented in this paper is easily scalable to cover larger areas, by incorporating, in the design, the information about the Wi-Fi access points in the corresponding areas. It is also one of the most cost-effective technologies compared to the other ones discussed in Section III. The only cost of the system is the human capital needed to design and implement the system; assuming the environment already has existing Wi-Fi infrastructure, which could be assumed to be the case in most indoor environments these days.

Google Indoor Maps [15] is also found to be a promising tool that provides an API and development system for indoor positioning and navigation. This could potentially be a great resource for developing an INS prototype in the future. The API utilizes Wi-Fi infrastructure for the positioning. The advantage of this tool is that it makes it possible to access Google services such as routing and location estimation, which would not need to be created by the developer from scratch. It would also help reduce the complexity of the app development to handle features such as zooming and scrolling. Unfortunately, at the time or writing this paper, Google was mainly focused on supporting developments for large shopping malls, sports complexes, and airports; not small projects like that of ours.

VI. CONCLUSIONS

This paper presented a low-cost and effective Wi-Fi based indoor positioning system and utilized the A* algorithm for path planning. The localization method employed in the project relies on the availability of good Wi-Fi infrastructure in the indoor environment. The WKNN algorithm deployed for this purpose is found to be a relatively simpler approach compared to the alternative methods, such as the fingerprinting technique. Through the project prototype it has been demonstrated that the system has acceptable level of accuracy for many common indoor navigational applications.

A typical application scenario that is a target use case for this project is providing navigational aid for university campus visitors. Currently, our university trains and employs students to serve as tour guides for visiting prospective students and their parents. Campus visits are offered on a daily basis and they are often conducted by the tour guides to take the visitors to specific locations of their interest (labs, classrooms, offices, etc.) One of the main benefits of this INS app would be to offer visitors the opportunity to lead their own tour by freely exploring the campus at their own pace. In future work, the map coverage of the app could be expanded and additional multimedia material could be created to provide audio-visual guide for the visitors to enhance their campus visit experience.

REFERENCES

[1] J. Golenbiewski, G. Tewolde, and J. Kwon, "Evaluation of Indoor Positioning Technologies for Prototyping at Kettering University," in *The 2018 IEEE Internation Conference on Electro/Information Technology (EIT 2018)*.

[2] "iBeacon," [Online]. Available: https://developer.apple.com/ibeacon/.

[3] "Google Beacon Platform," [Online]. Available: https://developers.google.com/beacons/.

[4] "How to do accurate indoor positioning with Bluetooth beacons?," [Online]. Available: https://proximi.io/accurate-indoor-positioning-bluetooth-beacons/.

[5] O. Pathak, P. Palaskar, R. Palkar, and M. Tawari, "Wi-Fi Indoor Positioning System Based on RSSI Measurements from Wi-Fi Access Points --A Tri-lateration Approach," *International Journal of Scientific & Engineering Research,* vol. 5, no. 4, pp. 1234-1238, 2014.

[6] C. Rizos, A. G. Dempster, B. Li, and J. Salter, "Indoor positioning techniques based on wireless LAN," in *LAN, First IEEE International Conference on Wireless Broadband and Ultra Wideband Communications,* 2007.

[7] B. Shin, J. H. Lee, T. Lee, and H. S. Kim, "Enhanced weighted K-nearest neighbor algorithm for indoor Wi-Fi positioning systems," in *The 8th International Conference on Computing Technology and Information Management (NCM and ICNIT)*, 2012.

[8] IndoorAtlas, "IndoorAtlas," [Online]. Available: https://www.indooratlas.com/.

[9] "Indoor Positioning - Perferct Light, Precise Location," [Online]. Available: http://www.lighting.philips.com/main/systems/themes/led-based-indoor-positioning.

[10] Pozyx, "How does ultra-wideband work?," [Online]. Available: https://www.pozyx.io/pozyx-academy/how-does-uwb-work.

[11] A. Suprem, V. Deep and T. Elarabi, "Orientation and Displacement Detection for Smartphone Device Based IMUs," *IEEE Access,* vol. 5, pp. 987-997, 2017.

[12] J. Golenbiewski and G. Tewolde, "Implementation of an Indoor Positioning System using the WKNN Algorithm," in *The 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC 2019)*, 2019.

[13] D. Ferguson, M. Likhachev, and A. Stentz, "A Guide to Heuristic-based Path Planning," in *The International Conference on Automated Planning & Scheduling (ICAPS)*, 2005.

[14] A* algorithms. [Online]. Available: http://theory.stanford.edu/~amitp/GameProgramming/.

[15] "Go Inside with Indoor Maps," [Online]. Available: https://www.google.com/maps/about/partners/indoormaps/.

Compact Six-Port Network Using Lange Coupler For Sub-6 GHz Fifth Generation Communication

Phule Anish Sunil
Department of Electrical Engineering
Indian Institute of Technology
Palakkad
Kerala, India
aphule01@gmail.com

Soumik Dey
Department of Electrical Engineering
Indian Institute of Technology
Palakkad
Kerala, India
soumikdey1990@gmail.com

Sukomal Dey
Department of Electrical Engineering
Indian Institute of Technology
Palakkad
Kerala, India
sdey28@iitpkd.ac.in

***Abstract*—This paper presents a wide band six port network (SPN) with two input ports and four out ports. The designed network consists of conventional wideband Lange couplers and a Wilkinson power divider having features of out of band harmonic suppression over a wide frequency. Operation band of the SPN extends from 2.3 - 4.0 GHz covering IEEE 802.11 WLAN band at 2.4 GHz and sub-6 GHz band from 3.3 - 3.8 GHz for 5G application. A good matching at the input ports are achieved with return loss of better than 20 dB for the entire frequency range with minimum isolation between the ports is 12 dB. At frequency 3.5 GHz the output powers at the four ports are -5.56 dB, -5.47 dB, -7.07 dB, -7.16 dB with input at port-1 and the powers are -6.05 dB, -7.8 dB, -5.02 dB, -6.56 dB with input at port-2. Differential phase between two adjacent output ports are ± 90 degree.**

Keywords—six port network, Lange coupler, Wilkinson power divider, broadband, quadrature phase.

I. INTRODUCTION

In modern microwave system passive components play an important role due to simple configuration and low losses. Growth in planar printed fabrication technologies like microwave monolithic integrated circuit (MMIC), low temperature co-fired ceramic (LTCC) is one of the key reason behind the high performance and miniaturization of the passive components. Passive six port network (SPN) is gaining attention as an alternative method of complex reflection coefficient (Γ) measurement at microwave frequency [1]–[2]. Unlike the traditional heterodyne scheme for measuring Γ of a load, six port reflectometer uses the power measured at its four port to determine the complex value of Γ. Each of the four ratios of power readings P_k for the device under test (DUT) to a reference power P_{ref} represents a circle in a complex Γ plane and the unknown reflection coefficient is determined from the common intersection point of the three circles [3]. Hoer for the first time presented SPN as sensing device for measuring voltage, current, power at four ports along with complex impedance and phase information [4]. In vector network analyzer (VNA), SPN are used as compact device with fully integrated ultra wideband application due to the accuracy in measuring the magnitude and phase of complex signal by heterodyning receiver technology. SPN also emerges as microwave receiver in recent years [5]–[6]. Conventional heterodyne receiver restrains from achieving compactness and design on a single chip because of filters at both RF and IF section. In [5] SPN as a wide band millimeter wave receiver for quadrature phase shift keying (QPSK) communication has been proposed. The direct conversion method of SPN will reduce the circuit complexity and allowing a higher level of circuit integration than conventional receiver. Recently SPN composed of five traditional power dividers and a hybrid coupler has implemented to formulate a measurement setup for extracting noise wave correlation matrix of two port DUT [7]. In [8] Ibrahim et al. proposed a SPN comprises of only Wilkinson power dividers of two kinds–in phase and quadrature phase at the outputs. The presented approach replaces the traditional SPN consists of coupler and power divider. But the transmission coefficient at the four ports are only −11±2 dB due to use of 20 chip resistors. That will incur more complexity in fabrication and constraint it from low power applications. Same group also reported SPN as QPSK modulator with symbol rate of 400 Msymbols/s over an octave bandwidth of 66 % [9]. Other application of SPN as software defined radio (SDR) receiver has shown in [10]–[14]. Mathematical model of reconfigurable dual band SDR receiver architecture has developed in recent years [13]. The SDR uses six port correlator circuit to down convert the two received signal with dual carrier frequencies 2.5 GHz and 3.0 GHz. In [15], coupled coplanar waveguide (CPW) coupler has adopted to make the SPN with maximum amplitude imbalance 2.8 dB and phase error of ± 10° between frequency of 1.5 GHz to 2.5 GHz. A SPN with two isolated ports, two coupled ports and remaining two of input and direct port are reported in [16], where all ports are on a single layers and the network has implemented using CPW interdigitated structure. Edge-coupled coupled line microstrip coupler has integrated with 3 dB Wilkinson's power divider to build a six port reflectometer [17]. Recently substrate integrated waveguide (SIW) based SPN has implemented at millimeter wave frequency [18]–[19]. SIW designs will reduce the radiation losses from the junction and bend in the circuit particularly at high frequency. A dual layer SPN using SIW at the center frequency of 60 GHz has presented in [18].

In this paper a compact SPN has designed with a wide operation bandwidth from 2.3–4 GHz. The microstrip based single layer six port circuit consists of three Lange couplers and a 3 dB modified Wilkinson's power divider. Power divider presented in [20] having features of band pass characteristics and ultra wide band harmonic suppression has utilized in this work for better out of band rejection. The adopted technique in the designed SPN will help in achieving better signal to noise ratio for application in high power microwave receiver. Numerical simulation of the designed SPN has performed using high frequency structure simulator (HFSS). Reflection coefficient at the input ports are less than −20 dB. Maximum amplitude imbalance at the output ports

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

are 1.8 dB and phase error of ± 2.7°. The proposed SPN can be used as a microwave reflectometer by attaching power detectors at the output ports or as a receiver at sub-6 GHz covering the 5G frequency for mobile communication. Rest of the paper is organized as follows–design procedures of quadrature phase Lange coupler and the harmonic suppressed power divider are presented in section II. Geometry and simulated scattering results of the SPN combining the power divider and coupler is shown in section III. Finally, paper is concluded in section IV.

II. Design of Coupler and Power Divider

A. Geometry of Lange Coupler

In various microwave applications coupler is considered as one of the key component. As a broad band coupler, Lange coupler is most popular due to have an octave or more bandwidth. Wide bandwidth has obtained due to multiple coupled line based design configuration [21]. Lines are tightly coupled with each other due to small spacing between them. Also fringing fields from both edges of a line contribute to increase the coupling. Fig. 1 shows the geometry of the single layer Lange coupler. Taconic TLY-5 with relative permittivity ε_r=2.2 and loss tangent tanδ=0.0009 is used as a dielectric substrate. Height of the substrate is h=0.51 mm. Copper with conductivity $\sigma=5.8\times10^7$ S/m and thickness of 35 micron is used as conducting material to print the metal layer of the Lange coupler. Even and odd mode impedances in the interdigitated lines are Z_{0e} and Z_{0o} respectively [22]

$$Z_{0e} = Z_{0o}\frac{(c+q)}{(k-1)(1-c)} \tag{1}$$

$$Z_{0o} = Z_0\left(\frac{1-c}{1+c}\right)^{1/2}\frac{(k-1)(1+q)}{(c+q)+(k-1)(1-c)} \tag{2}$$

k= even no. of strips=4 and c= coupling coefficient and Z_0 is the terminating impedance=50 Ω. q is related to c and k [23]

$$q = \left[c^2+(1-c^2)(k-1)^2\right]^{1/2} \tag{3}$$

The required shape ratios w_c/h and g_c/h are found using (4)

$$\left(\frac{w_c}{h}\right)_{se/so} = \frac{8}{p}\left[\frac{p.(7+4/\varepsilon_r)}{11}+\frac{(1+1/\varepsilon_r)}{0.81}\right]^{1/2} \tag{4}$$

P is determined from (5)

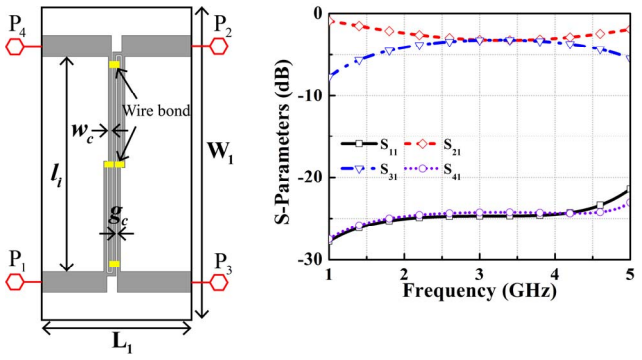
$$p = \left[\exp\left\{\frac{Z_{0o,0e}}{84.8}\left(\varepsilon_r+1\right)^{1/2}\right\}-1\right] \tag{5}$$

Synthesis equation for W/H is changes to (6) [21]

$$\left(\frac{w_c}{h}\right)_{so} = \frac{2}{\pi}\cosh^{-1}\left[\frac{(g+1).f-2}{(g-1)}\right]+r.\cosh^{-1}\left[\frac{\cosh^{-1}\left\{\frac{(g+1).f}{2}+\frac{(g-1)}{2}\right\}}{\cosh^{-1}g}\right] \tag{6}$$

Where $f = \cosh\left(\frac{\pi}{2}\left(\frac{w_c}{h}\right)_{se}\right)$ and $r = \frac{8}{\pi}(\varepsilon_r+2)$ for $\varepsilon_r \le 6$.

Equation (6) has been solved using different values of variable g. Value of g is always positive for a given ratios of $(w_c/h)_{so}$ and $(w_c/h)_{se}$. Fig. 2 in [22] can be used as a stating values of g for a given dielectric constant ε_r, after that iterative method has applied for a given tolerance. Final value of g is used to obtain the coupler dimensions w_c and g_c from (7) and (8) [22]



(a) (b)

Fig. 1. Configuration of Lange coupler (a) prototype geometry (b) simulated results of scattering parameter.

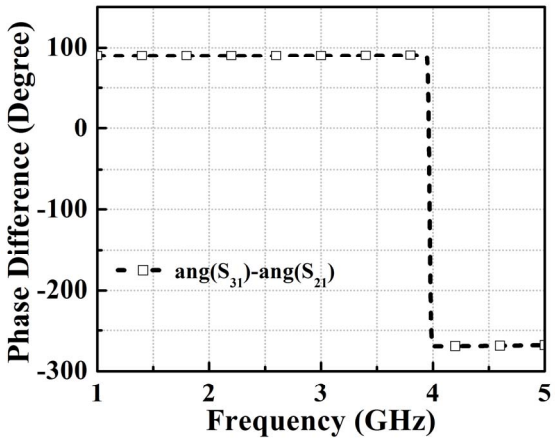


Fig. 2. Differential phase between the coupled and direct ports of designed Lange coupler.

$$w_c/h = \frac{1}{\pi}(\cosh^{-1}H - \cosh^{-1}g) \tag{7}$$

$$g_c/h = \frac{2}{\pi}(\cosh^{-1}g) \tag{8}$$

Where, $H = \frac{(g+1).f}{2}+\frac{(g-1)}{2}$ (9)

Length of the coupler strip $l_i = \frac{c}{4f\sqrt{\varepsilon_r}}$ (10)

Using (1) and (2) impedances are obtained as Z_{0e}=176.2 Ω and Z_{0o}=52.61 Ω. At the center frequency of 3.5 GHz using design equations (4)–(10), physical parameters of the Lange coupler are found to be w_c=0.24 mm, g_c=0.1 mm and l_i=14.45 mm. Ansys high frequency structure simulator (HFSS) is used to model the Lange coupler and all the rest of the components in this paper. Geometrical configuration and the simulated results of the coupler are shown in Fig. 1. For this four ports coupled line based coupler, P_1 to P_4 can be identified as input, direct, coupled and isolated ports. Even number of coupled line strips are used in the coupler geometry that are connected with each other using gold plated bonding wires. Optimized parameters are l_i=16.25 mm, L_1=11.26 mm, W_1=23.58 mm.

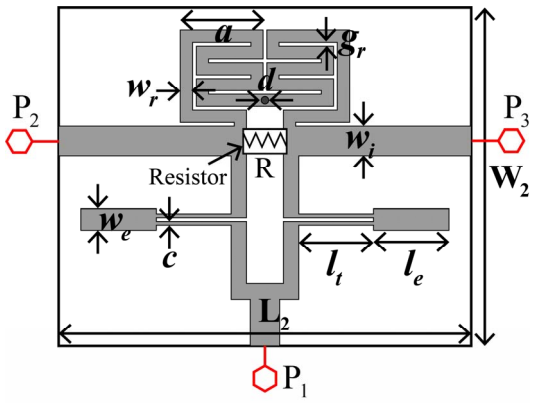
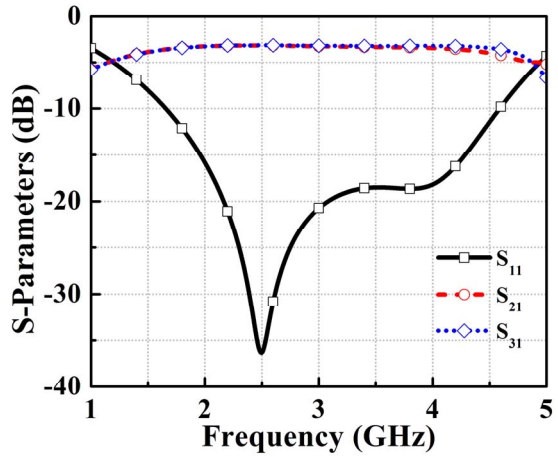
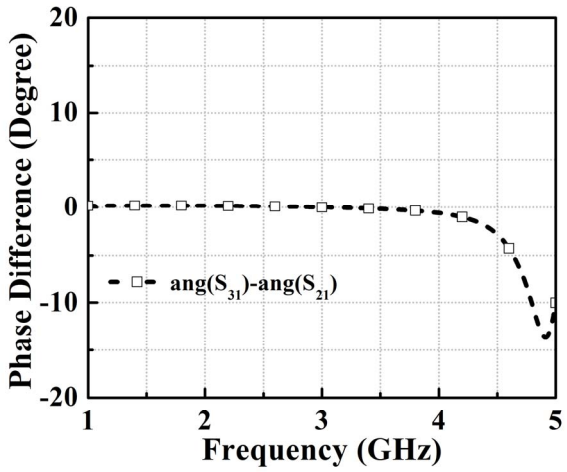


Fig. 3. Geometry of Wilkinson's power divider. Optimized design parameters are: L_2=21.965, W_2=18, l_t=3.98, l_e=4.02, w_e=1.14, w_i=1.565, w_r=0.66, g_r=0.2, a=4.41, d=0.4, c=0.2. All units are in mm.



(a)



(b)

Fig. 4. Simulated frequency response of WPD (a) S-parameter magnitude (b) phase difference between the output ports.

matching at the input port is better than -24 dB within the operating range of 2.3–4 GHz and isolation is constant throughout the band and average values is -24.28 dB. Amplitude imbalance between the direct and coupled ports is less than 0.1 dB with in the operating range except at lower band edge where the difference is 0.99 dB between S_{21} and S_{31}. Phase difference between the two output ports is 90° as depicted in Fig. 2. Wide operational band of the designed

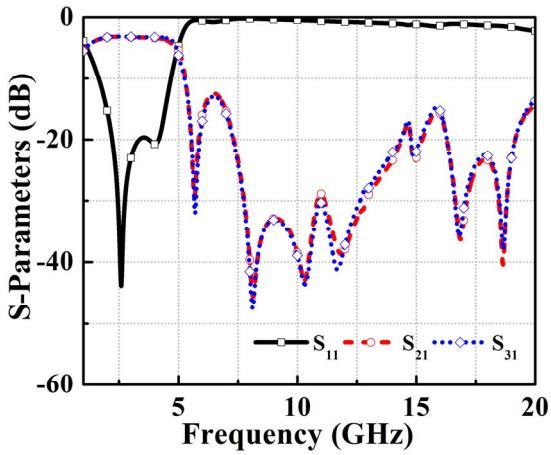


Fig. 5. S-parameter magnitudes of wide harmonic suppress WPD.

Lange coupler will assist in broadband performance of the proposed six port network as will be discusses in section III.

B. Harmonic Suppressed Wilkinson Power Divider

Power divider or power combiner is an essential part of six port network. Wilkinson power divider (WPD) are mostly used in microwave circuits with equal or arbitrary power division. Signals at the output ports can be either in-phase or in quadrature phase. In this paper an in-phase WPD has been designed with a broad frequency range. To improve the out of band rejection and reduce the losses due to mode dispersion higher order harmonic of the designed WPD has also suppressed using the designed technique as reported in [20]. Prototype configuration of the WPD is presented in Fig. 3. The overall size of WPD is $L_2 \times W_2$=21.965 mm × 18 mm or 0.256 λ_0 × 0.21 λ_0 at the center frequency of 3.5 GHz. Harmonic suppression unit in the WPD comprises of U shaped open ended horizontal stubs symmetrically attached with the two output branches. Two arms of U-shaped stub have electrical length of $\pi/2$. Coupling between them provides six transmission zeroes outside the operating region [20]. Band pass nature of the WPD is obtained by using shorted meander stubs loaded in the output ports. Presence of the shorted stub blocks the dc and the low frequency component due to additional transmission zero near to the lower cutoff frequency. A chip register having value of R=100 Ω is attached between the two output port to improve the isolation between the output ports and absorb the reflected power due to impedance mismatch at the outputs [21]. Reflection coefficient at the input and the output transmission characteristics have been simulated and the results are shown in Fig. 4. Within the operating frequency range outputs are close to -3 dB with average values of S_{21} and S_{31} are respectively -3.186 dB and -3.143 dB. A zero-degree phase difference is maintained over a wide band between the output signals due to symmetric configuration. Fig. 5 presents that harmonic suppression characteristics of the designed WPD over a wide frequency range has obtained as suggested in [20]. Stop band rejection of better than -10 dB is achieved for frequency more than 5.3 GHz with at least six transmission zeroes outside the frequency band. This helps to restrain the harmonics from 2nd to 7th order within an acceptable limit and that assists to improve the performance of six port network in microwave receiver as discussed in the following section.

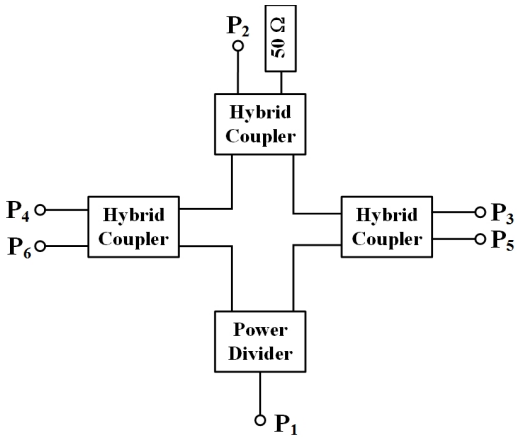


Fig. 6. Block diagram of the designed six port network.

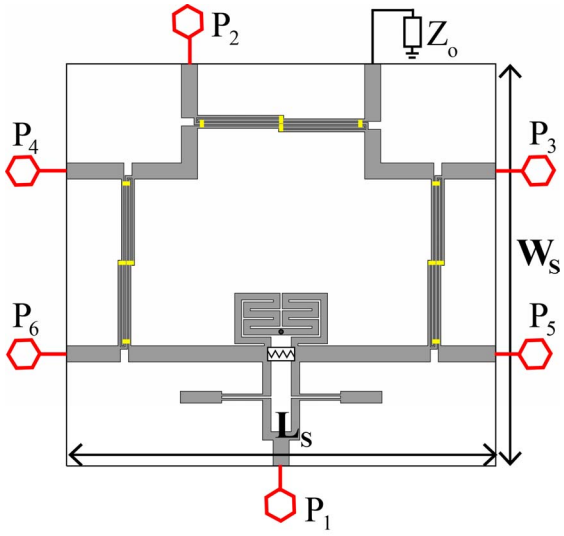
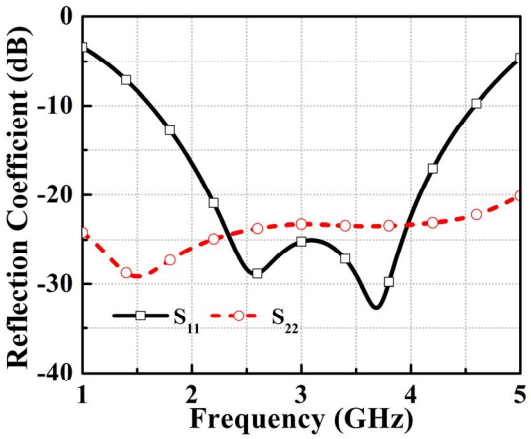
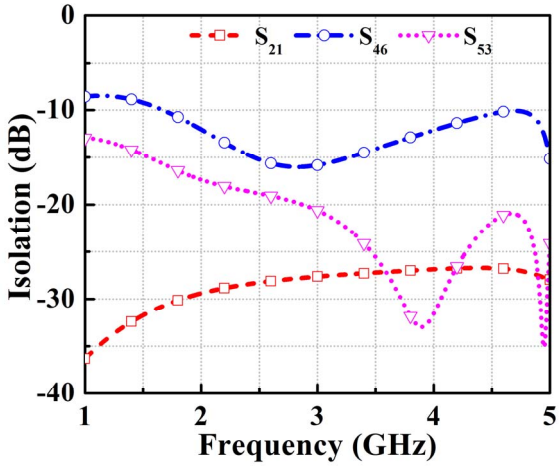


Fig. 7. Geometry of the designed six port network. Lateral dimensions are L_S=41.72 mm and W_S=39.235 mm.



(a)



(b)

Fig. 8. Simulated (a) reflection coefficient at the input ports and (b) isolation between the ports.

III. Design of Proposed Six Port Network

Block diagram of the proposed six port network has shown in Fig. 6. It comprises of three broadband Lange coupler and one harmonic suppressed WPD. SPN can be designed to determine the complex reflection coefficient in a microwave receiver. In the designed SPN ports P_1 and P_2 are the input ports and P_3 to P_4 are used as outputs. One of the ports of a coupler is matched terminated with 50 Ω resistance to make the total number of active ports in the circuit to be six. Assuming complex incident signals at the input ports 1and 2 be x_1 and x_2 the S-parameters at four output ports are S_3 to S_6 which are determined to be (11) [8]

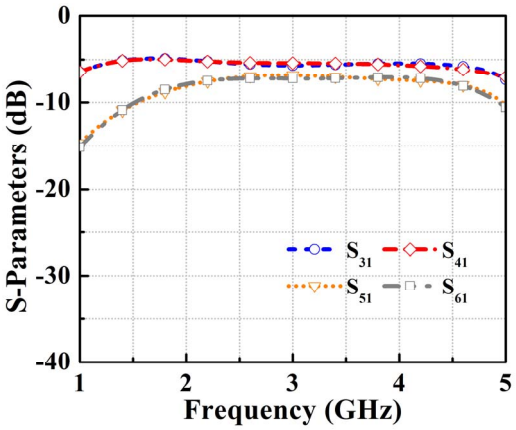
$$S_3 = -j\frac{x_1}{2}\left[\frac{x_2}{x_1} + j\right] \tag{11a}$$

$$S_4 = -j\frac{x_1}{2}\left[\frac{x_2}{x_1} + 1\right] \tag{11b}$$

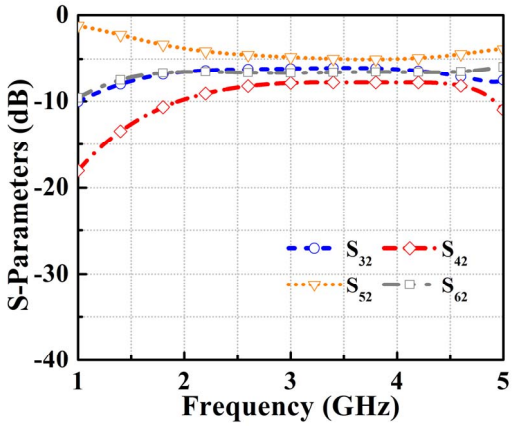
$$S_5 = \frac{x_1}{2}\left[\frac{x_2}{x_1} - j\right] \tag{11c}$$

$$S_6 = \frac{x_1}{2}\left[\frac{x_2}{x_1} - 1\right] \tag{11d}$$

Fig. 7 depicts the prototype of the designed SPN which is made on same Taconic TLY-5 substrate with thickness 0.51 mm. The total size of the network is $L_S \times W_S$= 41.72 mm × 39.235 mm or 0.487 $\lambda_0 \times 0.458\ \lambda_0$. Conventional topology that has adopted in the design of SPN helps in reducing the network complexity and achieving broadband performance. Since both input and output ports of the network are in same plane, the designed SPN can be easily integrate with different components in microwave receiver. The configuration results in low loss and wide band matching with high isolation between the ports. Magnitudes and phase of the scattering parameters for the designed SPN are obtained using full wave simulation and the results are illustrated in Fig. 8 to Fig. 10. Impedance matching within the operating frequency range 2.3–4 GHz is better than -23 dB for both the input ports as shown in Fig. 8(a). Decoupling between the input and output ports are well below −15 dB except for S_{46} which is slightly more near to the band edges (see Fig. 8(b)). Fig. 9 shows S-parameter magnitudes at the output ports are close to the theoretical value of −6 dB. Maximum amplitude imbalances at 3.5 GHz are 1.16 dB and 1.79 dB for input at port 1 and 2 respectively. Differential phase between the output ports are found out near to ± 90° as given in Fig. 10. Maximum phase errors at 3.5 GHz are determined to be 2.67° and 1.23° with inputs at port-1 and 2 respectively. Simulated magnitude and



(a)



(b)

Fig. 9. Simulated output S-parameters with input excitation at (a) port-1 (b) port-2.

phase responses validate the performance of the designed SPN suitable for future 5G application.

IV. Conclusion

A broadband compact six port network consists of three Lange couplers and a power divider has presented in this paper. The total electrical size of the network is only 0.487 λ_0 × 0.458 λ_0 × 0.0059 λ_0. The designed network shows a wide band matching over the operating frequency of 2.3–4 GHz. The insertion loss from network is less than 1.8 dB with magnitudes of output S parameters are close to -6 dB and phase differences between adjacent output ports are close to ±90°. Input excitations at ports 1 and 2 cause maximum average magnitude imbalances of 1.19 dB and 1.98 dB respectively over the operating frequency range. Phase errors for the presented SPN is less than 4.7° inside the operating range. The designed SPN can be used as sub-6 GHz microwave receiver and also as QPSK modulator for wireless digital communication. This will be done in the future work.

Acknowledgment

The authors would like to express their profound gratitude to the Director of Indian Institute of Technology Palakkad for setting up Central Instrumentational Facilities. This work is supported by Science and Engineering Research Board, Govt. of India under the project no: ECR/2018/002258.

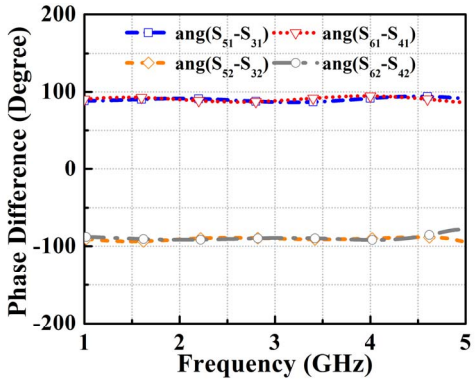


Fig. 10. Phase characteristics of the designed six port network.

References

[1] G. F. Engen, "The Six-Port Reflectometer: An Alternative Network Analyzer," *IEEE Trans. Microw. Theory Techn.*, vol. 25, no. 12, pp. 1075-1080, Dec. 1977.

[2] G. F. Engen, "An Improved Circuit for Implementing the Six-Port Technique of Microwave Measurements," *IEEE Trans. Microw. Theory Techn.*, vol. 25, no. 12, pp. 1080-1083, Dec. 1977.

[3] J. J. Yao and S. P. Yeo, "Six-Port Reflectometer Based on Modified Hybrid Couplers," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 2, pp. 493-498, Feb. 2008.

[4] C. A. Hoer, "The Six-Port Coupler: A New Approach to Measuring Voltage, Current, Power, Impedance, and Phase," *IEEE Trans. Instrum. Meas.*, vol. 21, no. 4, pp. 466-470, Nov. 1972.

[5] S. O. Tatu, E. Moldovan, K. Wu and R. G. Bosisio, "A new direct millimeter-wave six-port receiver," *IEEE Trans. Microw. Theory Techn.*, vol. 49, no. 12, pp. 2517-2522, Dec. 2001.

[6] T. Hentschel, "The six-port as a communications receiver," *IEEE Trans. Microw. Theory Techn.*, vol. 53, no. 3, pp. 1039-1047, Mar. 2005.

[7] A. Ahmed and K. Yeom, "An Extraction of Two-Port Noise Parameters From Measured Noise Powers Using an Extended Six-Port Network," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 10, pp. 2423-2434, Oct. 2014.

[8] S. Z. Ibrahim, A. Abbosh and M. Bialkowski, "Design of wideband six-port network formed by in-phase and quadrature Wilkinson dividers," *IET Microw. Antennas Propag.*, vol. 6, no. 11, pp. 1215-1220, Aug. 2012.

[9] S. Z. Ibrahim, A. M. Abbosh and M. A. Antoniades, "Direct quadrature phase shift keying modulation using compact wideband six-port networks," *IET Microw. Antennas Propag.*, vol. 6, no. 8, pp. 854-861, Jun. 2012.

[10] C. de la Morena-Álvarez-Palencia, M. Burgos-García and D. Rodríguez-Aparicio, "Three octave six-port network for a broadband software radio receiver," *in Proc. 40th Eur. Microw. Conf.*, Paris, France, Sep. 2010, pp. 1110-1113.

[11] C. de la Morena-Álvarez-Palencia and M. Burgos-Garcia, "Four-Octave Six-Port Receiver and its Calibration for Broadband Communications and Software Defined Radios," *Progr. Electromagn. Res.*, Vol. 116, 1-21, 2011.

[12] C. de la Morena-Álvarez-Palencia, M. Burgos-Garcia, and J. Gismero-Menoyo, "Miniaturized 0.3-6 GHz LTCC Six-Port Receiver for Software Defined Radio," *Progr. Electromagn. Res.*, Vol. 142, 591-613, 2013.

[13] A. O. Olopade, A. Hasan and M. Helaoui, "Concurrent Dual-Band Six-Port Receiver for Multi-Standard and Software Defined Radio Applications," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 12, pp. 4252-4261, Dec. 2013.

[14] A. Hasan, M. Helaoui, N. Boulejfen, and F. M. Ghannouchi, "Six-port technology for mimo and cognitive radio receiver applications," *in Proc. IEEE Topical Conf. on Wireless Sensors and Sensor Networks* (WiSNet), San-Diego, USA, Jun. 2015, pp. 17–19.

[15] S. Qayyum and R. Negra, "Compact six-port junction using broadside-coupled CPW couplers for a broadband six-port receiver," *in Proc. 47th Eur. Microw. Conf.*, Nuremberg, Germerny, Oct. 2017, pp. 347-350.

[16] Z. Wu, X. Hu and F. Xu, "A six-port network based on CPW interdigitated structure," *in Asia-Pacific Microw. Conf.*, Nanjing, China, Dec. 2015, pp. 1-3.

[17] D. Ghosh and G. Kumar, "Six-Port Reflectometer Using Edge-Coupled Microstrip Couplers," *IEEE Microw. Wireless Compon. Lett.*, vol. 27, no. 3, pp. 245-247, Mar. 2017.

[18] O. Kramer, T. Djerafi and K. Wu, "Dual-layered substrate-integrated waveguide six-port with wideband double-stub phase shifter," *IET Microw. Antennas Propag.*, vol. 6, no. 15, pp. 1704-1709, Dec. 2012.

[19] Xinyu Xu, R. G. Bosisio and Ke Wu, "A new six-port junction based on substrate integrated waveguide technology," *IEEE Trans. Microw. Theory Techn.*, vol. 53, no. 7, pp. 2267-2273, Jul. 2005.

[20] Y. Wang, X. Zhang, F. Liu and J. Lee, "A Compact Bandpass Wilkinson Power Divider With Ultra-Wide Band Harmonic Suppression," *IEEE Microw. Wireless Compon. Lett.*, vol. 27, no. 10, pp. 888-890, Oct. 2017.

[21] D. M. Pozar, Microwave Engineering, 4th ed. New York: Wiley, 2012, pp. 359–362, ch. 7.

[22] R. M. Osmani, "Synthesis of Lange Couplers," *IEEE Trans. Microw. Theory Techn.*, vol. 29, no. 2, pp. 168-170, Feb. 1981.

[23] P. K. Shrivastava, S. K. Koul and M. P. Abegaonkar, "Miniaturized 3-dB Micromachined Broadband Lange Coupler On Alumina Substrate," *2018 IEEE MTT-S Int. Microw. RF Conf.*, Kolkata, India, Nov. 2018, pp. 1-3.

Frequency Tuning Characteristics & Bandwidth Enhancement for Circular Microstrip Antenna by Integrating L Shaped DGS

[1]Srijita Chakraborty, [1]Debangana Dutta, [1]Pradipta Ghosh, [2]Neelanjana Giri,[2]Nipun Agarwal, [1]Malay Gangopadhyay, [2]Mrinmoy Chakraborty
[1]Institute of Engineering &Management, Kolkata, India
[2]Dr. B. C. Roy Engineering, College, Durgapur, India
Corresponding Email Id: srijita@iemcal.com

***Abstract*— In the ground surface a symmetrical defective ground structure (i.e. DGS) in L form was implemented for which circular microstrip antenna was resonant at second WLAN band , i.e. 5.8GHz. The geometry of the DGS was optimized to achieve lower application frequency bands viz. GSM band, Bluetooth band ,WiMax band and first WLAN respectively. As a result, the frequency tunability for the proposed antenna is observed in the application frequency bands with optimum bandwidth and gain. In order to achieve improved bandwidth at GSM and WiMax bands, dual L-shaped DGS is integrated into the ground plane of the proposed antenna.**

Keywords—Compact antenna, Circular Microstrip Antenna, Defected Ground Structure, L Shaped, Bandwidth Enhancement

I. INTRODUCTION

Extensive research on wireless microstrip antennas with characteristics such as less weight, low profile, cost efficient, portable and ease of fabrication which can be utilized for practical communications.

Meticulous efforts are being made to design and analyze highly miniaturized microstrip antennas with enhanced bandwidth characteristics. Various papers are reported in which compactness of microstrip antenna has been researched by etching symmetric slots, geometric slits, defective ground structures etc which results in a perturbation effect in the antenna[1,2]. Thus facilitating the antenna to be operated at a lower application frequency bands[3,4].

In the proposition, a symmetrical defective ground structure (i.e. DGS) in L shape has been implemented in the ground surface of a circular microstrip antenna, resonant at second WLAN application band i.e. 5.8GHz. The implementation of the DGS in the ground surface modifies the shielded current distribution. Then by the integration of L shaped DGS, the resonant frequency is reduced to 1.8GHz i.e. GSM frequency band. Therefore higher degree of compactness in size of the microstrip antenna is achieved. The length of the L shaped DGS is further altered to make the circular microstrip antenna resonant at 2.4GHz i.e. Bluetooth frequency band, 3.5GHz i.e. WiMax frequency band and 5.2GHz i.e. first WLAN frequency band respectively. Also by integrating dual DGS in the ground plane of the antenna, bandwidth enhancement has been achieved in GSM and WiMax application bands. Thus the proposed circular microstrip antennas can be favorably used in practical wireless applications.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

II. DESIGN PRINCIPLES

The detailed structural geometry of the proposed circular microstrip antenna is given in Fig.1. The dielectric substrate named FR4_epoxy with dielectric constant value of 4.4 and dielectric loss tangent value of 0.002 is employed to design the proposed circular microtsrip antenna and electromagnetic simulator tool Zeland IE3D is used to simulate the proposed microstrip antenna. In the proposed work, the circular microstrip antenna is designed so that it is resonant at second WLAN band i.e. 5.8 GHz. As DGS is installed into the ground plane, a frequency change from the WLAN band i.e. 5.8GHz to the GSM band i.e.1.8GHz is observed. The radius of the circular patch is 7mm and the feed point is placed at (0,3). For the afore mentioned dimensions, the proposed circular microstrip antenna is resonant at second WLAN band i.e. 5.8 GHz. The symmetric L- shaped DGS, etched in the ground surface is given in Fig. 2. The total length of the L- shaped DGS is almost 38.6mm with the width of 0.5mm. Due to integration of DGS at the ground surface, the proposed antenna is found to be resonant at the GSM band i.e. 1.8GHz. Furthermore the length of the L- shaped DGS is altered accordingly so that the resonant frequency is shifted from 1.8GHz to 2.4GHz, 3.5GHz and 5.2GHz respectively. From Fig. 3 it can be observed that the length of the L- shaped DGS is 29.1 mm, due to which the microstrip antenna is resonating at the Bluetooth frequency band i.e. 2.4GHz. Similarly from Fig. 4 and Fig. 5, it is justified that as length of the DGS is changed to 19.4mm and 10.1mm, the microstrip antenna becomes resonant at the WiMax frequency band, i.e. 3.5GHz and first WLAN frequency band i.e. 5.2GHz. Therefore it can be observed that the microstrip antenna is found to be resonant at higher application frequency bands as the length of the L-shaped DGS is decreased, which explains the tunable feature of microstrip antenna. Further another L shaped DGS of thickness 1mm is incorporated in the ground surface of the circular antenna as shown in Fig. 6 and Fig 7. which resulted in bandwidth enhancement of the frequency bands for GSM application band and WiMax application bands respectively.

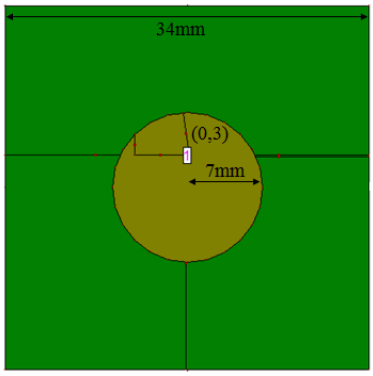


Fig.1. Front view of circular microstrip antenna , operational at 5.8GHz

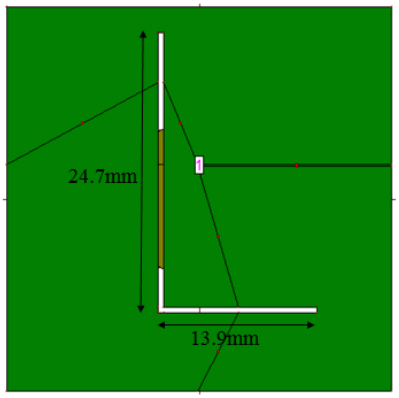


Fig.2. Rear view of DGS integrated circular microstrip antenna , operational at 1.8GHz

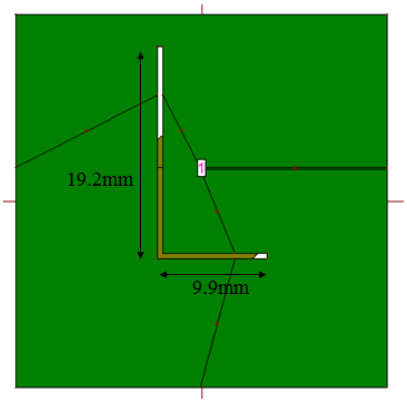


Fig.3. Rear view of DGS integrated circular microstrip antenna , operational at 2.4GHz

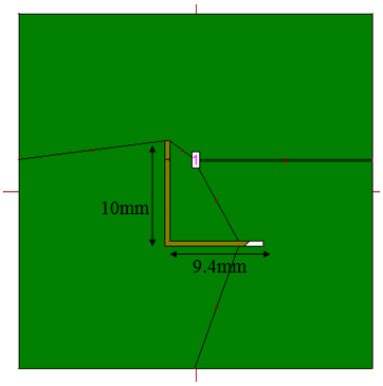


Fig.4. Rear view of DGS integrated circular microstrip antenna , operational at 3.5GHz

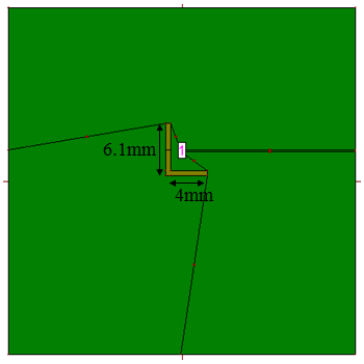


Fig.5. Rear view of DGS integrated circular microstrip antenna , operational at 5.2GHz

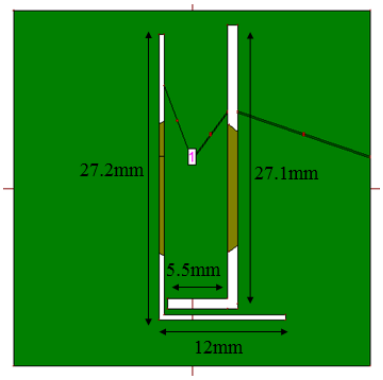


Fig.6. Rear view of dual DGS integrated circular microstrip antenna , operational at 1.8GHz with enhanced bandwidth characteristics

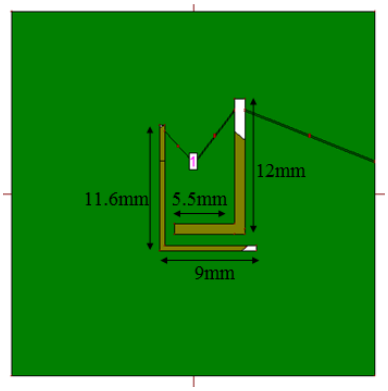


Fig.7. Rear view of dual DGS integrated circular microstrip antenna , operational at 3.5GHz with enhanced bandwidth characteristics

III. RESULTS

S_{11} versus frequency variation

The return loss versus frequency graph of the circular microstrip antenna with L shaped DGS and without L shaped DGS is shown in Fig.8 and Fig. 9 respectively. From the figures it is seen that with exclusion of DGS at the ground surface return loss at 5.8 GHz is -28dB and for the DGS integrated antenna the resonance is at 1.8GHz with -27dB return loss. Fig. 10 shows the variation of return loss versus frequency for the circular microstrip antenna with revised length of the DGS and it is seen that the antenna is resonating at Bluetooth frequency band i.e. 2.4GHz with -33dB return loss. Furthermore Fig. 11 and Fig.12 shows the variation of return loss versus frequency plot for the microstrip antenna, resonant at the WiMax frequency band, i.e. 3.5GHz and first WLAN frequency band i.e. 5.2GHz respectively. The return loss at resonance is obtained as -

24dB and -32dB subsequently. Also Fig. 13 and Fig. 14 demonstrates the return loss versus frequency plot for dual DGS integrated circular microstrip antenna for 1.8GHz, i.e. GSM band and 3.5GHz, i.e. WiMax band respectively.

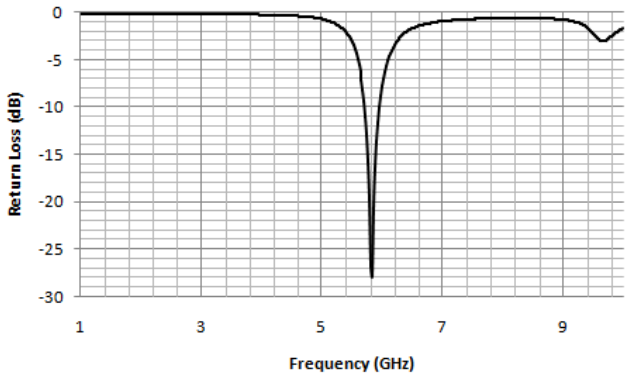


Fig.8. S_{11} versus frequency plot for circular microstrip antenna , operational at 5.8GHz

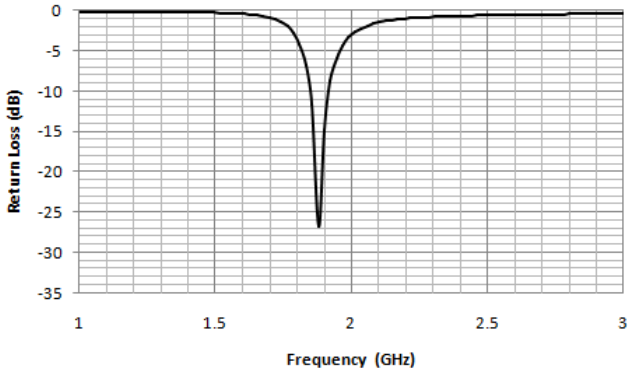


Fig.9. S_{11} versus frequency plot for DGS integrated circular microstrip antenna , operational at 1.8GHz

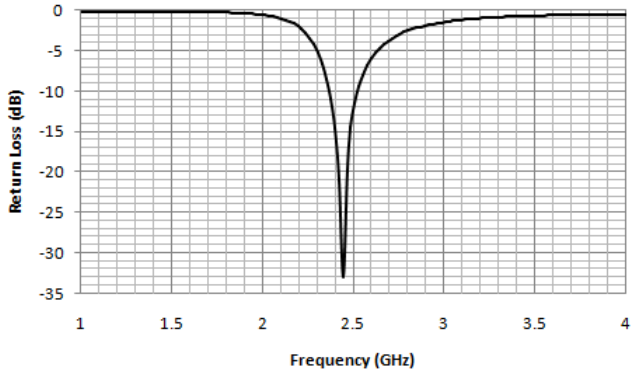


Fig.10. S_{11} versus frequency plot for DGS integrated circular microstrip antenna , operational at 2.4GHz

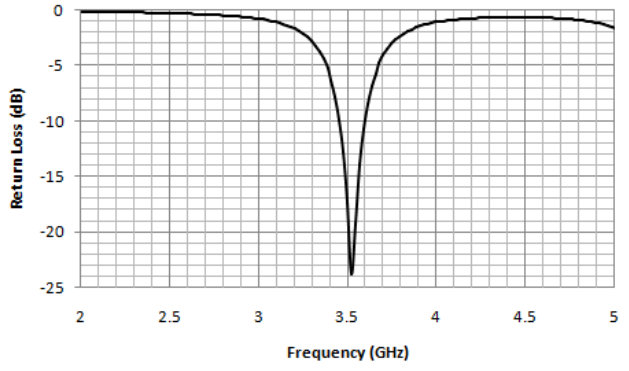


Fig.11. S_{11} versus frequency plot for DGS integrated circular microstrip antenna , operational at 3.5GHz

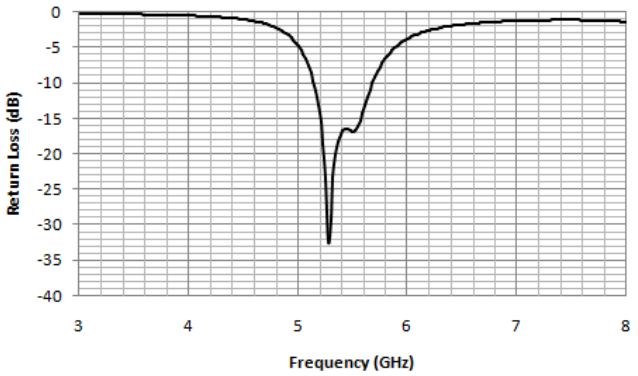


Fig.12. S_{11} versus frequency plot for DGS integrated circular microstrip antenna , operational at 5.2GHz

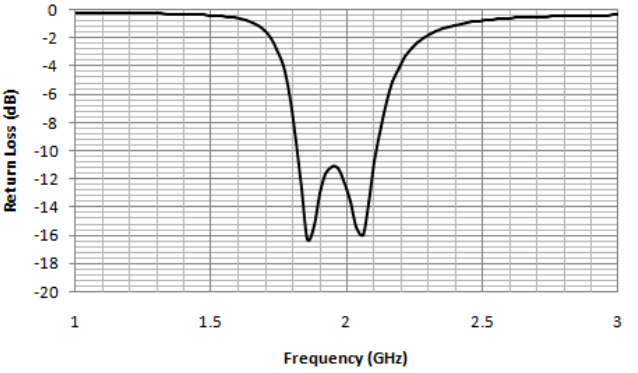


Fig.13. S_{11} versus frequency plot for dual DGS integrated circular microstrip antenna , operational at 1.8GHz with enhanced bandwidth characteristics

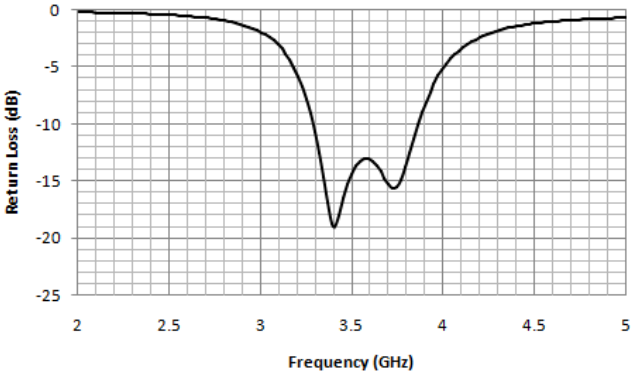


Fig.14. S_{11} versus frequency plot for dual DGS integrated circular microstrip antenna , operational at 3.5GHz with enhanced bandwidth characteristics

Impedance parameter versus frequency variation

At the resonance the resistive part of impedance in the impedance versus frequency graph for the microstrip antenna is the characteristics impedance of 50ohms. Similarly the net reactive part of impedance in the impedance versus frequency graph is 0ohm for optimal performance. Fig.15 shows the impedance versus frequency graph for the microstrip antenna without DGS and Fig.16 gives the same graph for DGS integrated microstrip antenna, resonant at 1.8GHz. Similarly Fig.17, Fig. 18. and Fig. 19 shows the impedance versus frequency plot for the microstrip antenna, resonant at 2.4GHz, 3.5GHz and 5.2GHz respectively. Also Fig. 20 and Fig. 21 demonstrates the impedance versus frequency plot for dual DGS integrated circular microstrip antenna for 1.8GHz and

3.5GHz respectively. It can be observed from figures that the resistive part of impedance and the net reactive part of impedance at resonance is approximately 50 ohms and 0 ohm respectively. Therefore at the resonating frequency the compact microstrip antenna is perfectly matched for optimum performance.

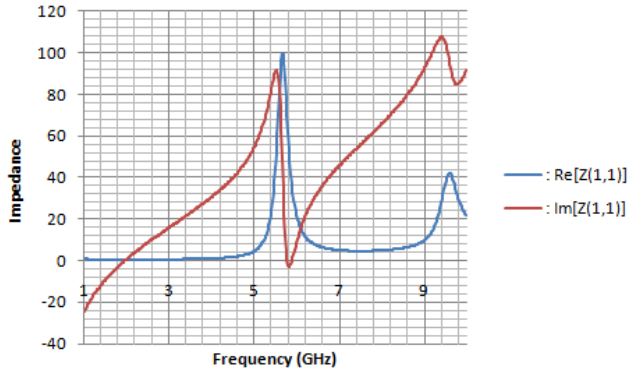


Fig.15. Z parameter frequency plot for circular microstrip antenna , operational at 5.8GHz

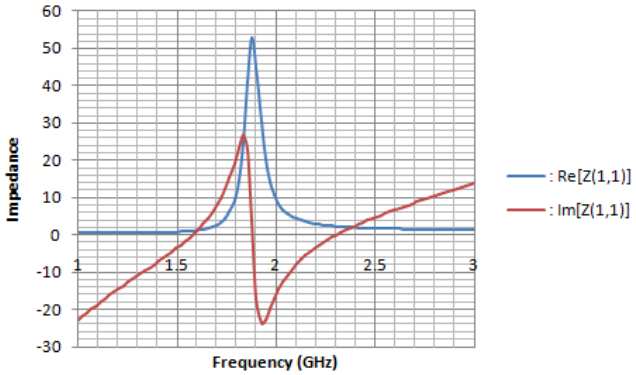


Fig.16. Z parameter versus frequency plot for DGS integrated circular microstrip antenna , operational at 1.8GHz

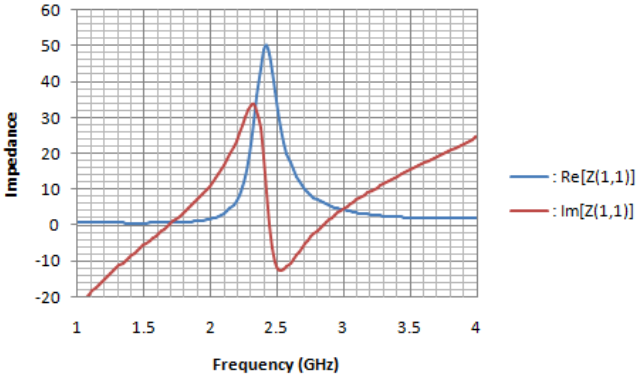


Fig.17. Z parameter versus frequency plot for DGS integrated circular microstrip antenna , operational at 2.4GHz

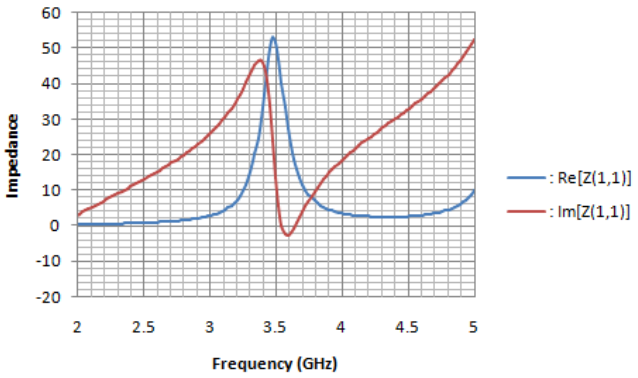


Fig.18. Z parameter versus frequency plot for DGS integrated circular microstrip antenna , operational at 3.5GHz

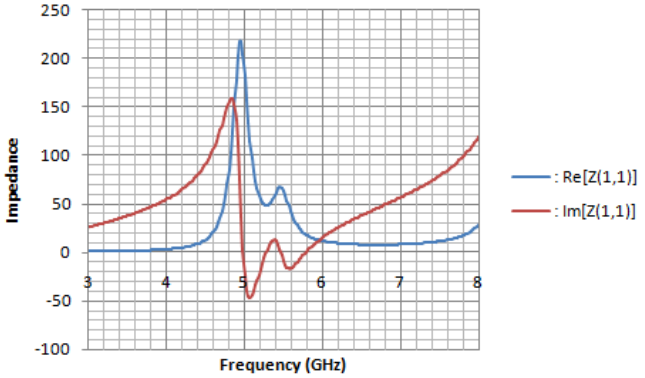


Fig.19. Z parameter versus frequency plot for DGS integrated circular microstrip antenna , operational at 5.2GHz

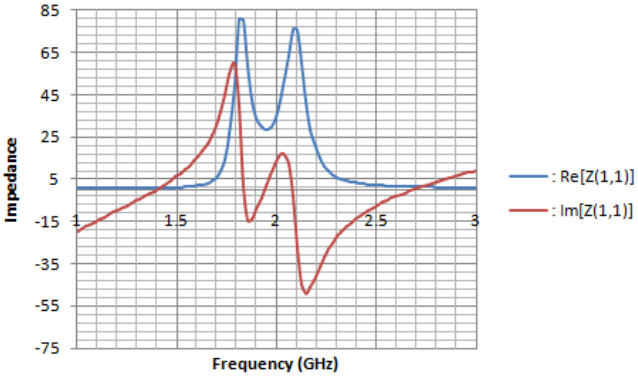


Fig.20. Z parameter versus frequency plot for dual DGS integrated circular microstrip antenna , operational at 1.8GHz with enhanced bandwidth characteristics

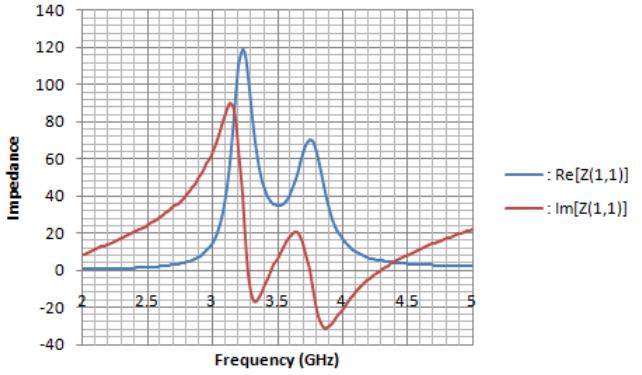


Fig.21. Z parameter versus frequency plot for dual DGS integrated circular microstrip antenna , operational at 3.5GHz with enhanced bandwidth characteristics

Radiation Pattern

The radiation plot of the microstrip patch antenna is orthogonal to the surface of the patch and so the elevation pattern when φ is 0 and 90 degrees are measured. The E-plane and H-plane radiation pattern at second WLAN frequency band (i.e. 5.8GHz) is shown in Fig.22 and it can be observed that the gain is obtained at resonant frequency is 4.4dBi. Fig.23 shows the E-plane and H plane radiation pattern at GSM band (i.e. 1.8 GHz) and the gain at 1.8GHz is obtained as 0.8dBi.Similarly Fig. 24, Fig. 25 and Fig. 26.shows the radiation pattern of the circular microstrip antenna resonating at 2.4GHz, 3.5GHz and 5.2 GHz respectively. From the figures it is observed that maximum gain as obtained are 1.7dBi, 2.1dBi and 3.9dBi respectively. Also Fig. 27 and Fig. 28 demonstrate the radiation pattern for dual DGS integrated circular microstrip antenna for 1.8GHz, i.e. GSM band and 3.5GHz, i.e. WiMax band respectively. Thus it can be observed that the gain at GSM and WiMax frequency band after inclusion of dual DGS is 0.8dBi and 2.21dBi. respectively.

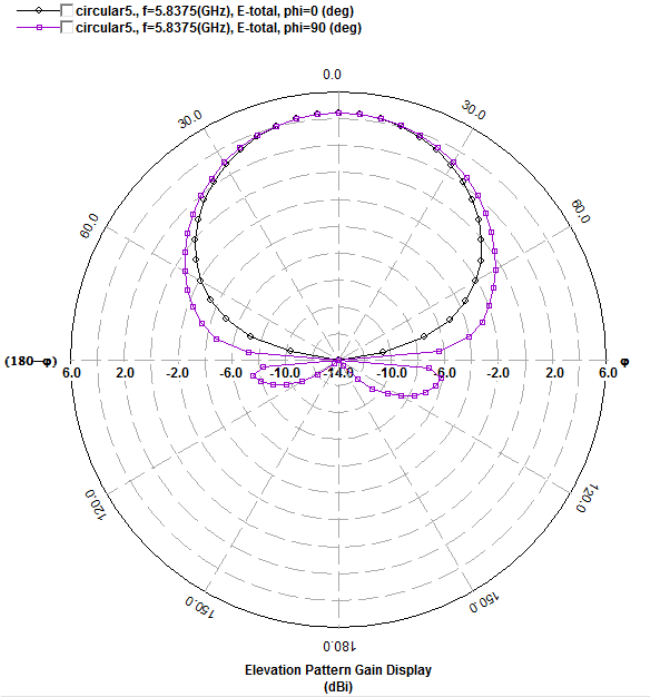


Fig.22. Radiation pattern for circular microstrip antenna , operational at 5.8GHz

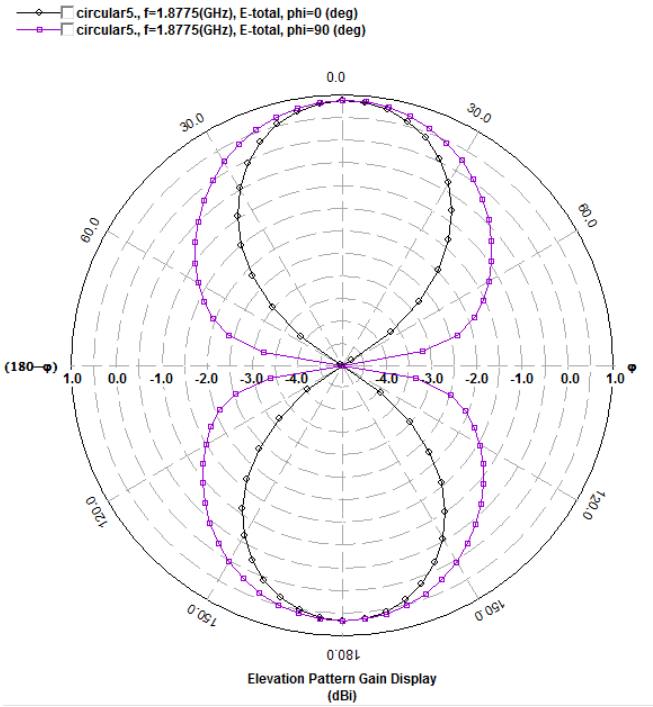


Fig.23. Radiation pattern for DGS integrated circular microstrip antenna , operational at 1.8GHz

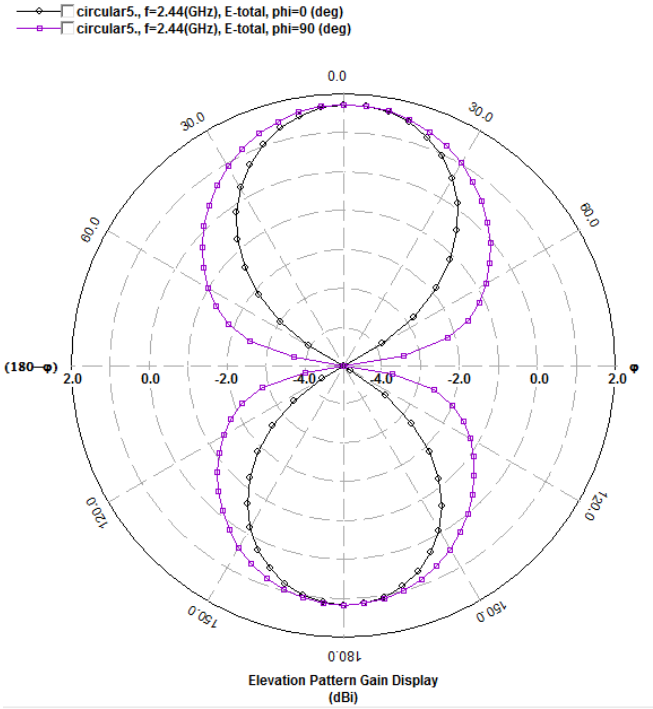


Fig.24. Radiation pattern for DGS integrated circular microstrip antenna , operational at 2.4GHz

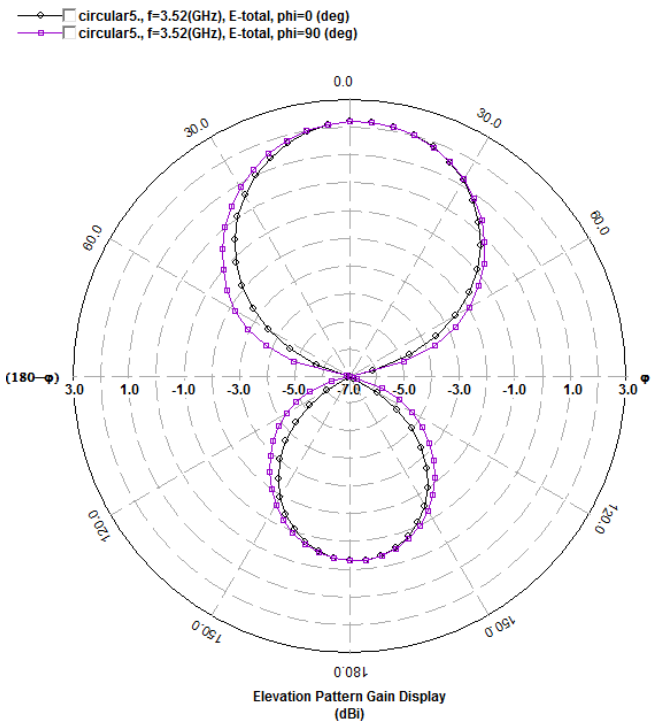


Fig.25. Radiation pattern for DGS integrated circular microstrip antenna , operational at 3.5GHz

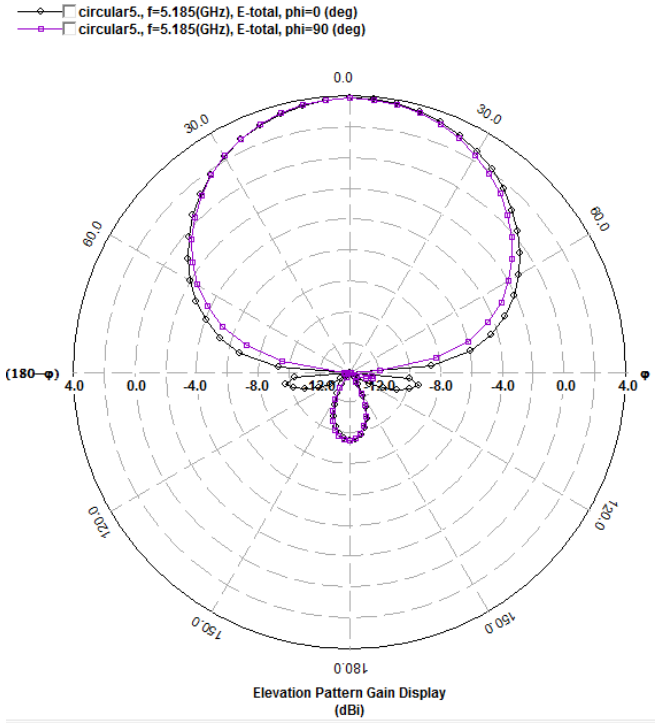


Fig.26. Radiation pattern for DGS integrated circular microstrip antenna , operational at 5.2GHz

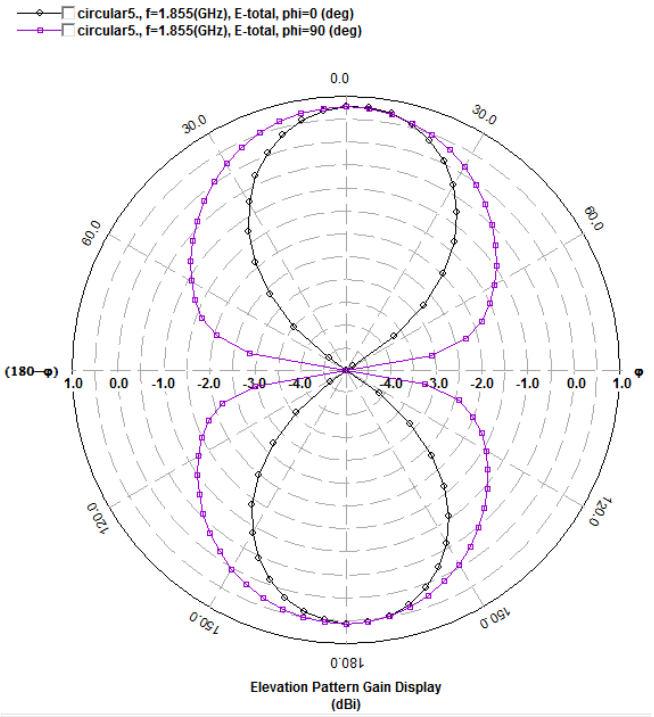


Fig.27. Radiation pattern for dual DGS integrated circular microstrip antenna , operational at 1.8GHz with enhanced bandwidth characteristics

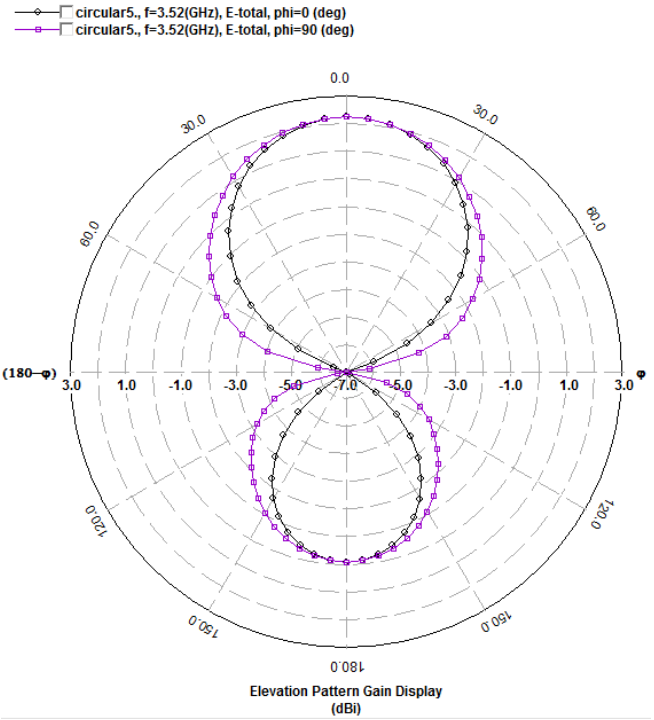


Fig.28. Radiation pattern for dual DGS integrated circular microstrip antenna , operational at 3.5GHz with enhanced bandwidth characteristics

IV. CONCLUSION

A L- shaped DGS implemented in a circular microstrip antenna results in appreciable compactness of about 91% and the shift of resonant frequency from the second WLAN band, 5.8GHz to GSM band, 1.8GHz, Bluetooth frequency band i.e. 2.4GHz, WiMax frequency band i.e. 3.5GHz and first WLAN band, 5.2GHz respectively. Also dual L shaped

DGS has been incorporated in the ground surface of the antenna to achieve bandwidth enhanced characteristics. This proves the tunability and bandwidth enhancement of the proposed circular microstrip antenna. Extremely high compactness and appreciable bandwidth in different application bands justify the optimal performance of the circular microstrip antenna. Thus the proposed micro strip antenna is suitable for implementation in practical wireless communications.

ACKNOWLEDGMENT

The proposed research work is done in Dr. B. C. Roy Engineering College, Durgapur, India and the authors would sincerely acknowledge the contribution of all concerned.

REFERENCES

[1] Mrinmoy Chakraborty, Biswarup Rana, P.P. Sarkar, Achintya Das,"Design and Analysis of a Compact circularMicrostrip Antenna with slots using Defective Ground Structure", Elsevier Science Direct Procedia Technology 4 (2012) 411 – 416

[2] Srijita Chakraborty, Mrinmoy Chakraborty, Srikanta Pal ,"Design and Analysis of a Compact circularMicrostrip Patch Antenna with Defected Ground Structure", Proceeding of 4th International Conference On Technical And Managerial Innovation in Computing And Communications in Industry And Academia, IEMCON2013,pp. 281-283

[3] S. Bhunia, D. Sarkar, S. Biswas,P. P. Sarkar, B.Gupta,and K. Yasumoto,"Reduced Size Small Dual and Multi-Frequency Microstrip Antennas",MICROWAVE AND OPTICAL TECHNOLOGY LETTERS / Vol. 50, No. 4, April 2008

[4] Srijita Chakraborty, Mrinmoy Chakraborty, Biswarup Rana, N.N.Pathak, Achintya Das "Design and Analysis of a Compact circular Microstrip Antenna with Slots and Defected Ground Structure", Proceedings of National Conference on Electronics and Communication Systems,2013,Section-3, Chapter 2, Page 150

Resonant Frequency Optimization of Cylindrical Liquid Antenna Using Particle Swarm Optimization Algorithm

Sutapa Ray*, Soham Kanti Bishnu, Agniva Chatterjee, Malay Gangopadhyay
Department of Electronics and Communication Engineering (ECE)
Institute of Engineering and Management
Kolkata, India.
*sutapa.ray@iemcal.com

***Abstract*— In recent years liquid antennas have evolved as a favourable option for several use. In this paper we have tried to present a optimized cylindrical liquid antenna using Particle Swarm Optimization (PSO) algorithm and the simulation of the said antenna was done using CST. Investigation shows that generated electric field line changes the effective radius of cylinder which in turn changes the resonant frequency. So good optimization technique for proper calculation of parameters was required which will consider both the electric field line effect and designed resonant frequency. In this paper only resonant frequency is optimized and PSO has been used to calculate radius (w) and height (L) of the liquid in order to get desired result. Experiment was carried out at different frequencies by varying the parameters such as the inner radius of the cylinder, height of the liquid, and the dielectric.**

Keywords*— *liquid antennas, cylindrical liquid antenna, Particle Swarm Optimization (PSO).

I. INTRODUCTION

The performance of antenna has a very vital role as it determines the data transmission speed and reliability of the system. So the modern communication requires the antenna to have multifunctionality and adaptability to dynamic environment in a wide range [1-9]. Several attempts have been made to enhance the reconfigurability of metallic antennas in the areas of radiation pattern, operating frequency and bandwidth. Low cost radio frequency (RF) switches does not perform well due to the problem of having metal strip biasing circuits which would be difficult to embed with antennas. Also, the high performance RF switches are very expensive [10-11]. Thus; studies have been made on dielectric resonators as radiating element. The dielectric resonator antenna is a kind of resonator fabricated using low-loss microwave dielectric materials. Extensive research is going on to study the effects of different parameters on DRA [13].

Liquids have high conformability, re-configurability, low radar cross section, transportability and have improved electromagnetic coupling. Due to several advantages water is used as dielectric in liquid antenna. Here Particle swarm optimization [14] have been used to design the liquid antenna at 4 GHz for Bluetooth utilization. Previously most of the liquid antennas worked in the range of in the frequency range from 0.5 GHz to 8.5 GHz.

II. DESING AND OPTIMIZATION OF CYLINDRICAL LIQUID ANTENNA

A. WATER ANTENNA CONFIGURATION

The main objective of this paper is to present a cylindrical liquid antenna which has been optimized by using the most popular optimization algorithm that is Particle Swarm Optimization (PSO) [14]. At first the proposed antenna has been designed using CST (Computer Simulation Technology) [13] considering the three parameters namely the radius of the liquid, height of the liquid, and the dielectricat 4GHz frequency. We varied the radius of the liquid keeping other parameters unchanged and its effect was observed on resonant frequency which is given below [15]

$$f_0 = \frac{19.972 \times\times 10^8}{2\pi a\sqrt{\epsilon_r}+2}\left[0.27 + 0.36\left(\frac{a}{2h}\right) + 0.02\left(\frac{a}{2h}\right)^2\right]$$

Where h, a, and $\epsilon_r$are the height, radius, and permittivity of a cylindrical dielectric, respectively, and $f_0$is the resonant frequency. The antenna is designed at 4 GHz for which the optimum inner radius=7mm, optimum height of the liquid =17 mm and the dielectric constant is taken as 65. The results show that the return loss is -20db at said frequency.

B. PARTICLE SWARM OPTIMIZATION

The Particle Swarm Optimization (PSO) Algorithm [16] is one the popular algorithm used.Thisoptimization process inspired from nature is based on social behavior of bird flocking and fish schooling.The basic idea of PSO is to exchange the information between particles of the swarm toexplore thesearch space in order to find an optimal solution tothe problem. In this algorithm, it is necessary to create random k particles, which form a swarm. All the variables in each particle must be optimized. Each variable must becreated within the specified range limits. The displacement of the particle is given by the sum ofthree vectors namely suffering inertia (*w)*, local best position (p_{best}) and global best position (g_{best}).

After the last iteration, the variables of the vector G_{best}correspond to theoptimum values of the vector **I** as described in Fig.1 [16]

978-1-7281-9615-2/20/$31.00©2020 IEEE

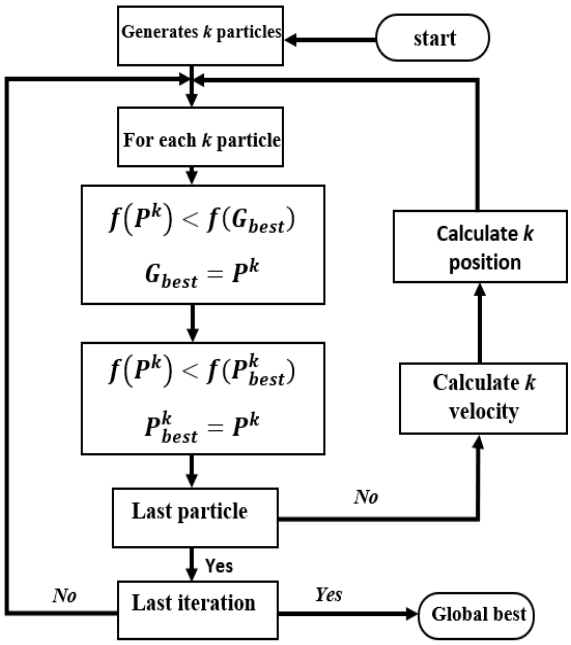


Fig. 1 Particle Swarm Optimization, flowchart

III. RESULTS

The study was done by using CST (Computer Simulation Technology) software to design the antenna as shown in Fig. 2. Table 1 and 2 shows the simulation result. The resonant frequency was also optimized .The return loss for different resonant frequencies were plotted as shown in Fig 3, 4 and 5 and the deviation of the optimum result and the resonant frequencieswasreported minimum.

The PSO optimizer parameters settings as follows:
Swarm size: 3
No of iterations: 20
Total no. of solver evaluations: 61
Weight: 1.0
Target: 4GHz

Table 1: The effects of radius on resonant frequencies

Sl. No	Optimum Height (mm)	Optimum Radius (mm)	Resonant Frequency (GHz)	Return Loss (dB)
1	17	5	4.19	-48.71
2	17	5.72	4.15	-46.6
3	17	6.06	4.09	-34.33
4	17	6.63	4.02	-23.04
5	17	7	4	-20.4
6	17	8.29	3.82	-19.69
7	17	8.31	3.8	-19.82
8	17	10.67	3.65	-37.65
9	17	13.26	3.64	-11.41
10	17	15.622	3.53	-14.16

From the simulation result as shown in Table 1, it shows that change in radius changes the resonant frequencies. As the proposed antenna was designed at 4 GHz thus by implementing the PSO technique the radius at designed frequency was successfully optimized.

The optimum results found are as follows:

Table 2: Optimum results

Parameter	Optimum Value
Inner Radius of Cylinder	7mm
Height	17mm
Return Loss	-20.4dB
Gain	1.43dB

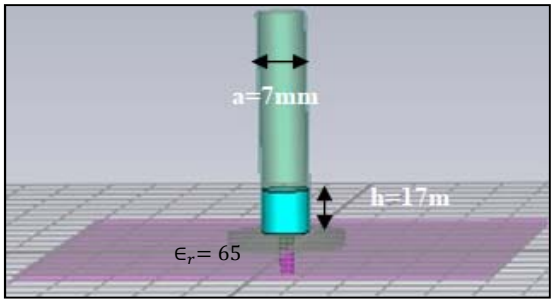


Fig. 2 Designed Cylindrical Liquid Antenna

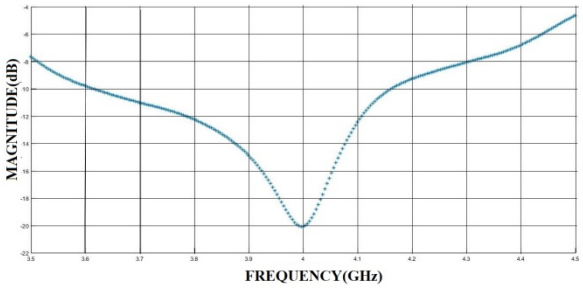


Fig. 3 Return loss plot for 4GHz resonant frequency

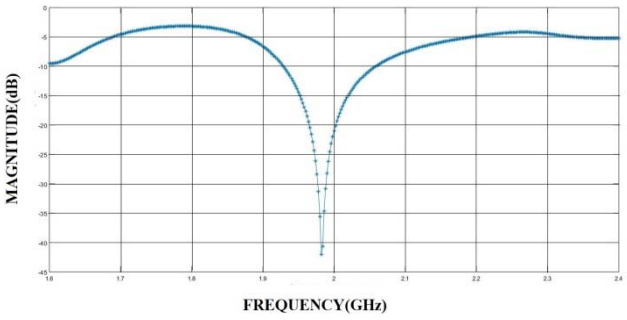


Fig. 4 Return loss plot for 2GHz resonant frequency.

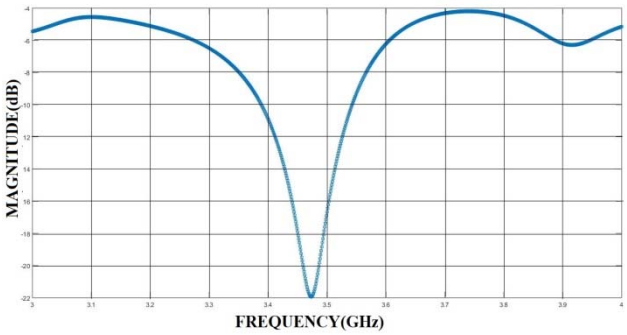


Fig. 5 Return loss plot for 3.5GHz resonant frequency.

IV. CONCLUSION

In this paper a thorough investigations on cylindrical liquid antenna has been studied carefully and it has been observed that the parametric changes have significant influence on the resonant frequency. So in case of liquid antenna optimization is obvious to restrict the deviation of the design frequency. From the results it may be concluded that by implementing PSO in our proposed antenna we have been able to obtain the optimum results at our desired frequency.

ACKNOWLEDGEMENT

We would like to express our special thanks of gratitude to Prof. Bhaskar Gupta and Dr.Kaushik Patra and for their continuous, motivation, and support.

REFERENCES

[1] J. Kiriazi, H. Ghali, H. Ragaie, and H. Haddara,"Reconfigurable Dual-band Dipole Antenna on Silicon usingSeries MEMS Switches," *in IEEE Antennas and PropagationSociety InternationalSymposium,* 2003, pp. 403–406.

[2] K. M. Ho, S. Member, and G. M. Rebeiz, "PolarizationDiversity and Frequency Agility*," IEEE Trans. AntennasPropag.,* vol. 62, no. 5, pp. 2398–2406, 2014.

[3] C. W. Jung, Y. J. Kim, Y. E. Kim, and F. De Flaviis, "Macromicrofrequency tuning antenna for reconfigurable wirelesscommunication systems," *Electron. Lett*., vol. 43, no. 4, p.201, 2007.

[4] M. R. Hamid, P. Gardner, P. S. Hall, and F. Ghanem,"Reconfigurable Vivaldi antenna with tunablestop bands," *in2011 International Workshop on Antenna Technology(iWAT)*, 2011, no. 1, pp. 54–57.

[5] X. Artiga, J. Perruisseau-carrier, P. Pardo-carrera, I. Llamasgarro,S. Member, and Z. Brito-brito,"Halved VivaldiAntenna With Reconfigurable Band Rejection," *IEEEAntennas Wirel. Propag. Lett.,* vol. 10, pp. 56–58, 2011.

[6] Jaehoon Kim and Y. Rahmat-Samii, "Low-profile antennasfor implantable medical devices:optimized designs forantennas/human interactions," *IEEE Antennas Propag. Soc.Symp*. 2004., vol. 2, p. 1331–1334 Vol.2, 2004.

[7] K.-F. Tong and J. Huang, "New proximity coupled feedingmethod for reconfigurable circularly polarized microstrip ringantennas," *IEEE Trans. Antennas Propag.,* vol. 56, no. 7,2008.

[8] P. J. B. Clarricoats and H. Zhou, "The design andperformance of a reconfigurable mesh reflector antenna," *in7th IEEE International Conference on Antennas andPropagation,* 1991, pp. 322–325.

[9] J. Chiao, M. Delisio, and R. Bank, "MEMS ReconfigurableVee Antenna*," in 1999 IEEE MTT-S InternationalMicrowave Symposium Digest,* 1999, pp. 1515–1518.

[10] Yue Li, Zhijun Zhang, Jianfeng Zheng, Zhenghe Feng, andM. F. Iskander, "Experimental Analysis of a WidebandPattern Diversity Antenna With Compact ReconfigurableCPW-to-Slotline Transition Feed," *IEEE Trans. AntennasPropag.,* vol. 59, no. 11, pp. 4222–4228, Nov. 2011.

[11] Guoan Wang, T. Polley, A. Hunt, and J. Papapolymerou, "A high performance tunable RF MEMS switch using bariumstrontium titanate (BST) dielectrics for reconfigurableantennas and phased arrays," *IEEE Antennas Wirel. Propag.Lett.,* vol. 4, pp. 217–220, 2005.

[12] Petosa, A., Ittipiboon, A., Antar, Y.M.M., et al.: 'Recent advances indielectric-resonator antenna technology', *IEEE Antennas Propag. Mag.,* 1998,40, (3), pp. 35–48

[13] Xing, Lei & Huang, Yi & Shen, Yao-chun & Aljaafreh, Saqer & Xu, Qian & Alrawashdeh, Rula. (2015). Further investigation on water antennas. *IET Microwaves Antennas & Propagation*. 9. 10.1049/iet-map.2014.0298.

[14] Gangopadhyaya, M., Mukherjee, P., & Gupta, B. (2010, December). Resonant frequency optimization of coaxially fed rectangular microstrip antenna using particle swarm optimization algorithm. In *2010 Annual IEEE India Conference (INDICON)* (pp. 1-3). IEEE.

[15] Mongia, R.K.; Bhartia, P. Dielectric resonator antennas—A review and general design relations for resonantfrequency and bandwidth. *Int. J. Microw. Millim. Wave Comput. Aided Eng.* 1994, 4, 230–247.

[16] E. R. Schlosser, S. M. Tolfo and M. V. T. Heckler, "Particle Swarm Optimization for antenna arrays synthesis," 2015 *SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC), Porto de Galinhas,* 2015, pp. 1-6, doi: 10.1109/IMOC.2015.7369120.

A Study on Scilab Free and Open Source Programming for Antenna Array Design

V. A. Sankar Ponnapalli
Department of Electronics and Communication Engineering
Sreyas Institute of Engineering and Technology
Hyderabad-500068, India
e-mail: sankar.p@sreyas.ac.in

Vinod Babu P
Department of Electronics and Communication Engineering
Rajiv Gandhi University of Knowledge Technologies
Nuzvid-521202, India.
e-mail: vinod@rguktn.ac.in

Abstract— **Usage of free and open source software has quietly increased in the scientific and academic community from the last few years. This study paper provides Scilab free and open source program example for the design of antenna arrays. A heptagonal antenna array has been considered and designed using Scilab programming. These results are compared with the same array antenna but implemented by the commercial software. Comparative analysis of array factor properies and how this software will be more opt for the academia has discussed in this paper.**

Keywords: *Scilab, Antenna array, Software, Program, Commercial software, Array factor*

I. INTRODUCTION

Open source software tools are the best solutions for the academia and research industries. Developing open source laboratory and research tools plays a vital role and they are need of the hour in this pandemic(COVID-19) situation across the globe. Various open source tools has been developing for different engineering and numerical applications. Especially commercial electromagnetic simulators imposes much financial burdens to academic institutions of the developing and underdeveloped countries, for such cases the open source simulating tools are absolutely a boon for the research communities and specially for individual students. In the way to find the solutions for electromagnetic problems using open source tools, this paper concentrated on a preliminary study on Scilab programming for the design of antenna arrays. Actually Scilab programming is entirely bestowed with a great programming language whose syntax is simple to employ and has all characteristics of a high level programming [1]. Owing to this advantage Scilab software is widely used in various fields of engineering. Merits of using Xcos and Scilab in the design of RC filter rather than simulink and matlab has discussed in [2], authors of this article appreciated that the Scilab syntax is available in their local language. Fiber optic communication tool box has been developed using Scilab. Few selected modules of this tool box are electrical filters, measuring instruments, signal sources, optical sources, components, measuring instruments, amplifiers, modulators, etc [3].

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Actual technical reports on the design of array antennas with Scilab and comparison study with commercial Electromagnetic software tools is very rare to find hence hardly we can get the literature survey. But in [4] antenna design using open source software tools has reported and this work concentrates on the open EMS, nec2 softwares for the design of conventional printed and yagi antennas. A detailed comparison between the execution and run timings of the open source and commercial codes were also discussed in [5]. In this paper a heptagonal planar array antenna has considered for demonstrating the Scilab programming. Antenna arrays are the only solution for the sophisticated communication systems like advanced mobile networks, satellite, radar systems and 5G Technologies. That's why theoretical analysis and synthesis of these antennas is very much essential for the establishment of proper array antenna systems [6, 7]. The usage of Scilab in the academia is also increasing day by day and the concepts like array antennas can easily understandable by the students because of open and online flexibility of the software. The statistics of Scilab usage in various domains in Indian academic institutions is given in the fig.1 [8, 9]. Next sections of this paper will explain the heptagonal antenna array and array factor equations. Sections III discuses the Scilab coding aspects, comparison study, results, and discussion. Finally conclusion draws in the last section of this paper.

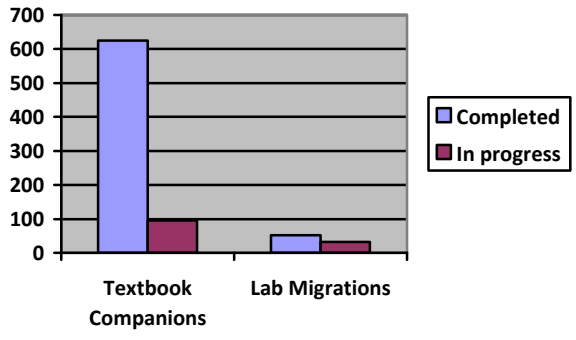


Fig. 1. Statistics of Scilab usage in India [9]

II. HEPTAGONAL ANTENNA ARRAY AND ARRAY FACTOR EQUATION

Basically various antenna arrays have reported in the literature for the wireless communication systems. Among these circular array antenna is one of the popular array for multiple Electromagnetic applications. In this work seven elements are considered in the circular surface, which means a heptagonal antenna array has been considered with uniform current distribution and maintaining the same distance between the antenna elements, as shown in fig. 2 [6, 7]. The corresponding array factor equation is presented in eq.(1).

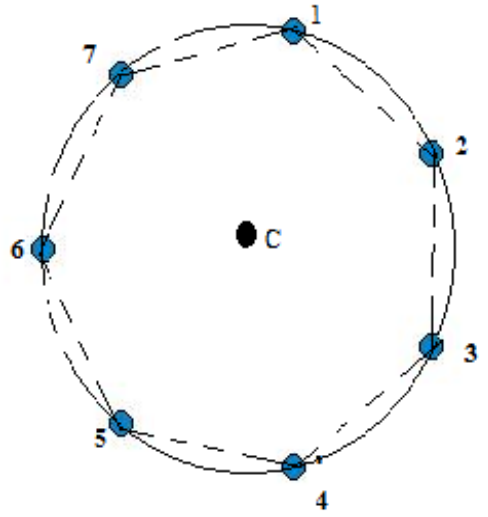


Fig. 2. Heptagonal antenna array.

$$A.F(\theta,\phi)=\sum_{n=1}^{7} I_n e^{jkr\sin\theta\cos(\phi-\phi_n)+\alpha_n} \quad (1)$$

where, I_n is the current distribution and here uniform distribution is considered and remaining elements represents the geometrical and array factor characteristics.

III. HEPTAGONAL ANTENNA ARRAY DESIGN WITH SCILAB AND COMMERCIAL PROGRAMMING

This section concentrated on the coding part of the considered array and comparison study of the Scilab and commercial software tools. The program of Scilab and resultant array factor graphs are depicted in the fig. 3.

A. Code Description

1, 2 – r and N are basic input parameters for the considered antenna array. Here radius (r) is 0.5λ and number of elements (N) is 7.

4, 5- Beam steering angles and here both angles considered as 0^o, and these angles can be varied according to the requirement.

6, 7, 11- Elevation and azimuth angles of the considered array antenna and array factor equation.

17- By this comment rectangular plot can be drawn as shown in fig. 3 (b) and with the polar syntax instruction, polar plot can be drawn as shown in fig. 3(c).

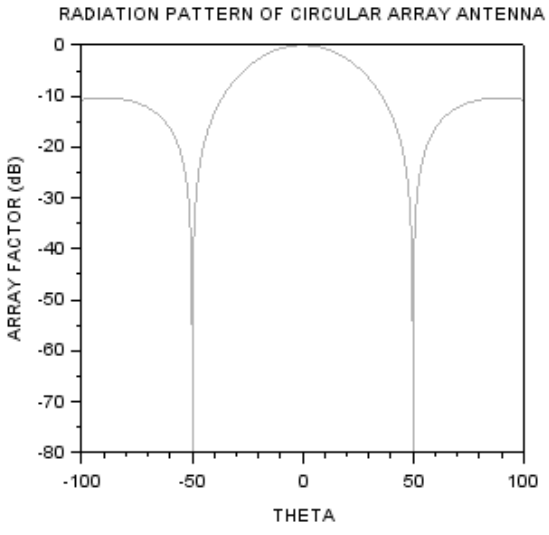
19- This instruction gives the specific axis boundaries of rectangular plot.

B. Results and Comparison Study

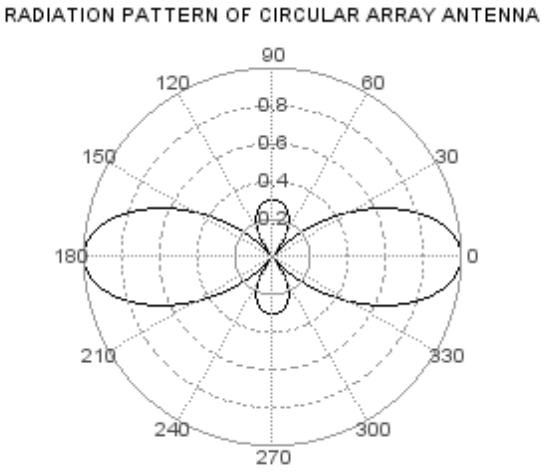
From the array factor plots as depicted in the fig. 3., the side lobe level of the array is -10 dB, with beamwidth of 41.8^o, side lobe level angle of 95.6^o has been observed. The array factor values of the commercial software are tabulated in the table. 1. There is no significant difference between the array factor values of the Scilab free and open source software and the commercial software. Owing to this reason it will be more useful for the students, researchers, and teachers. Further, It should be appreciable for academia to make use of this open source software particularly in this global pandemics shutdown situations like COVID-19.

```
1. r=input('RADIUS=');// IN WAVELENGHTS
2. N=input('NUMBER OF ANTENNA ELEMENTS=')
3. af=0;
4. bhio=0; //STEERING ANGLE
5. phio=0; // STEERING ANGLE
6. phi=%pi/2;
7. thi=-%pi:%pi/10000:%pi;
8. k=2*%pi
9.  for n=1:N // NUMBER OF ANTENNA ELEMENTS
10. phin=(n-1).*(2.*%pi/N)
 11. a=(exp(%i.*k.*r.*(((sin(thi)*cos(phi-phin)))-
(sin(thio)*cos(phio-phin)))));
 12.  af=a+af;
 13. end
14. afl=abs(af);
15. afl=afl/max(max((afl)));
16. CAA=20.*log10(afl);
17. plot((thi*57.3),(CAA));
18. h=gca();
19. h.data_bounds=[-90,-80;90,0];
```

(a)



(b)



(c)

Fig. 3. (a) Scilab coding of the considered antenna array (b) & (c) Array factor in rectangular and polar coordinates.

Table I. Result comparison of Scilab and commercial software of heptagonal antenna array.

	SLL (dB)	**Beam width (Deg.)**	**SLL Angle (Deg.)**
SCILAB Software	-10	41.8	95.6
Commercial Software	-10.3	41.8	88.9

IV. CONCLUSION

This article focused on the study of the Scilab open and free source software for the design of the heptagonal antenna array. It is observed from the study analysis, there is no significant difference in the Scilab and commercial software values of array factor. The run time of the both softwares is absolutely same. This study further extended into various levels like to analysis, and optimization of the conventional and advanced antenna systems. This software is more useful for the academia and Individual researchers.

REFERENCES

[1] Gomez, Claude. *Engineering and Scientific Computing with Scilab*. Springer Science & Business Media, 1999.

[2] Shakin, V. N., et al. "Comparison of Computer Modeling of RC Filter in Matlab and Scilab Environments." *2020 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, 2020.

[3] Jaworski, Marek, "Scilab open-source software for fiber optic communication systems simulation." *2015 17th International Conference on Transparent Optical Networks (ICTON)*, 2015.

[4] Fedeli, Alessandro, et al. "Open-Source Software for Electromagnetic Scattering Simulation: The Case of Antenna Design." *Electronics*, vol. 8, no. 12, pp. 1-31, 2019.

[5] Fedeli, Alessandro, et al. "Free and Open Source Software Codes for Antenna Design: Preliminary Numerical Experiments." *Electrical, Control and Communication Engineering*, vol. 15, no. 2, pp. 88-95, 2019.

[6] Sankar Ponnapalli, V. A., and P. V. Jayasree, "Heptagonal Fractal Antenna Array for Wireless Communications," *Microelectronics, Electromagnetics and Telecommunications, Lecture Notes in Electrical Engineering*, vol. 372, pp. 387-394, 2016.

[7] R. J. Mailloux, Phased Array Antenna Handbook, Artech House, Norwood, MA, 2005.

[8] https://scilab.in/. Accessed on 2020/02/03.

[9] https://fossee.in/fossee-stats, Accessed on 2020/05/15.

Implementation of Dielectric Resonator Integrated Microstrip Antenna for Improved Performance Characteristics

[1]Srijita Chakraborty, [1]Sayanti Dutta, [2]Apurba Sahu, [2]Aparna Kumari Singh, [1]Malay Gangopadhyay, [2]Mrinmoy Chakraborty
[1]Institute of Engineering &Management, Kolkata, India
[2]Dr. B. C. Roy Engineering, College, Durgapur, India
Corresponding Email Id: srijita@iemcal.com

Abstract- A dielectric resonator antenna is proposed in the given paper which results in increased gain and efficiency. Initially a simple rectangular microstrip antenna, 5.8GHz resonant was equipped with a 3.5dBi gain and a 60 per cent radiation output. A cylindrical structure is mounted to form a dielectric resonator on the simple microstrip antenna resulting in an improved 4.7dBi gain and an appreciable radiation output of 93 per cent. Finally, the idea of integrating dielectric resonator was introduced to further improve antenna gain. Thus two cylindrical holes are inserted over the microstrip antenna into the dielectric resonator, which achieves enhanced gain of 4.79dBi.

Keywords: **Microstrip antenna, DRA, Dielectric resonator antenna, Improved efficiency, Improved gain**

I. INTRODUCTION

Microstrip antennas have been created, analyzed and planned by researchers around the globe in the recent past. The reason for such extensive work in this field is due to the characteristic features of the microstrip antenna such as light weight , low profile, compact, with minimal cost and ease of manufacture. Despite the lucrative advantages of the antenna, certain inherent drawbacks such as narrow bandwidth , low performance , low gain, foreign radiation from feeds and junctions, low power handling capability, surface wave excitation etc. still cripple the antenna. Rigorous attempts have been made in this regard to remove the limitations of the microstrip antenna, so it can be used properly for practical wireless communication. At first, a rectangular microstrip-line feed microstrip antenna is built to be operational at 5.8GHz, with 3.5dBi gain and 60 percent radiation efficiency. The proposed structure achieves an improved gain of 4.7dBi and an appreciable radiation output of about 93 percent. Eventually, two cylindrical symmetric holes are inserted over the antenna into the cylindrical structure to further improve the gain of the dielectric resonator antenna, which provided an improved gain of 4.69dBi. Further analysis and studies in this regard are required to develop the implementation of dielectric resonator in microstrip antenna to improve the gain and radiation efficiency.

II. DESIGN PRINCIPLES

The detailed geometry of the microstrip rectangular antenna (antenna 1) is given in Fig.1. The substrate FR4 epoxy with dielectric constant 4.4 and the dielectric loss tangent of 0.02 are used for the antenna design, and the simulation is carried out using the electromagnetic simulation software HFSS. The microstrip antenna is 36 mm long and 28 mm wide, respectively. Although the length and width of the rectangular patch are respectively 13 mm and 11 mm. Likewise the length and width of the rectangular patch is 13 mm and 11 mm respectively. An inset feed of length 4.2m and width of 1.5mm is implemented and 12.7mm long microstrip line feed and 3mm wide is used to feed the rectangular microstrip antenna. The proposed antenna is found to be 5.8GHz resonant with a gain of 3.5dBi and a radiation efficiency of 60 per cent. To improve the gain and efficiency characteristics of the microstrip antenna, on top of the specified microstrip antenna (antenna 2) a cylindrical

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

structure with a length of 7.5mm and a radius of 4.5mm is installed, as shown in Fig. 2.

The proposed antenna is found to be resonant at 5.8GHz, with 3.5dBi gain and 60% radiation efficiency. To boost the gain and efficiency characteristics of the microstrip antenna, on top of the specified microstrip antenna (antenna 2) a cylindrical structure with a length of 7.5mm and a radius of 4.5mm is implemented, as shown in Fig.2. The material used for the cylindrical construction has a dielectric constant of 20 and tangent dielectric loss of 0. Gain of about 4.7dBi and sufficient radiation efficiency of about 93 per cent is achieved in the proposed dielectric resonator antenna.

Ultimately, in order to further maximize the gain of the proposed microstrip antenna, two 0.5 mm and 3 mm separate symmetrical holes are drilled through the cylindrical structure at the top of the microstrip antenna (antenna 3), as shown in Figure. 3.The proposed dielectric resonator antenna is observed to provide an enhanced gain of about 4.79dBi.

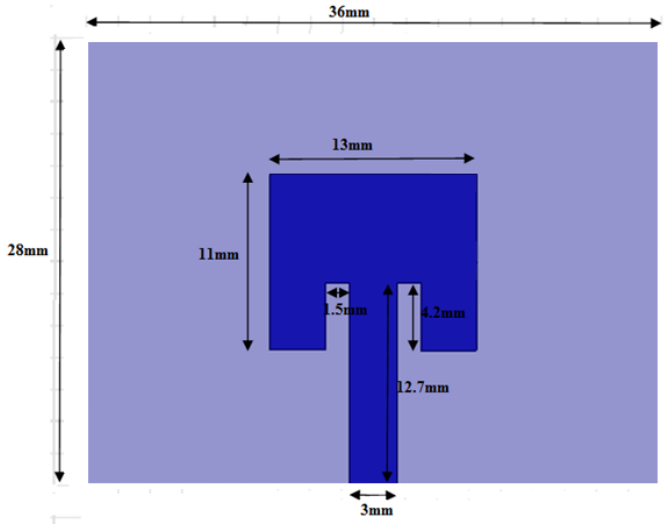


Fig. 1. Antenna 1- Top view of the proposed antenna, resonating at 5.8GHz

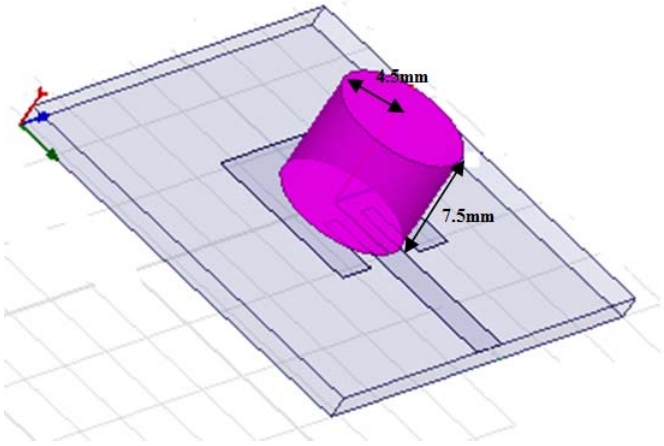


Fig. 2. Antenna 2-Angular view of the proposed antenna with cylindrical structure built on top(dielectric resonator antenna), resonating at 5.8GHz

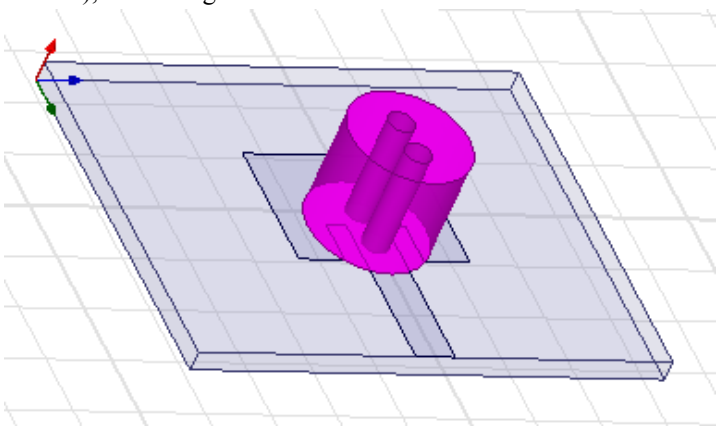


Fig. 3. Antenna 3-Angular view of the proposed DRA (dielectric resonator antenna) with cylindrical holes, resonating at 5.8GHz

III. RESULTS

S_{11} versus frequency graph

Fig.4 displays the S11 versus the frequency plot of the rectangular microstrip antenna, and from the figure it can be seen that the antenna is resonant at 5.8Ghz with a return loss of -28dB. Fig. 5 Shows the S11 versus the dielectric resonating antenna frequency plot (i.e. a microstrip antenna with a cylindrical structure mounted at the top) and it is observed that the antenna is resonant at 5.8GHz, with a return loss of -13dB. Finally Fig.6 displays the dielectric resonator antenna S11 versus frequency variation. It can be noted from the figure that the dielectric resonator antenna is resonant at 5.8GHz, with a return loss of -16dB.

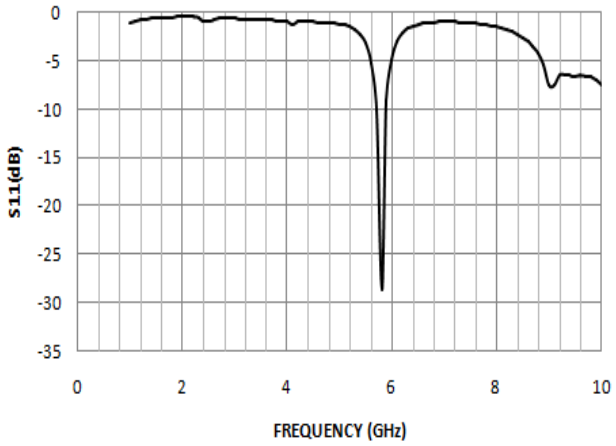


Fig. 4. Simulated S_{11} versus frequency of the proposed antenna , resonating at 5.8GHz

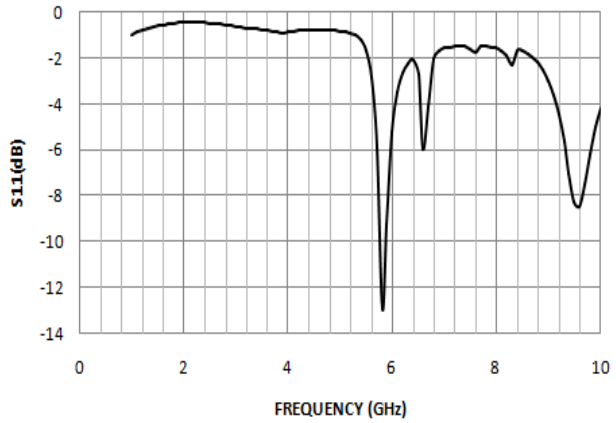


Fig. 5. Simulated S_{11} versus frequency of DRA (dielectric resonator antenna), resonating at 5.8GHz

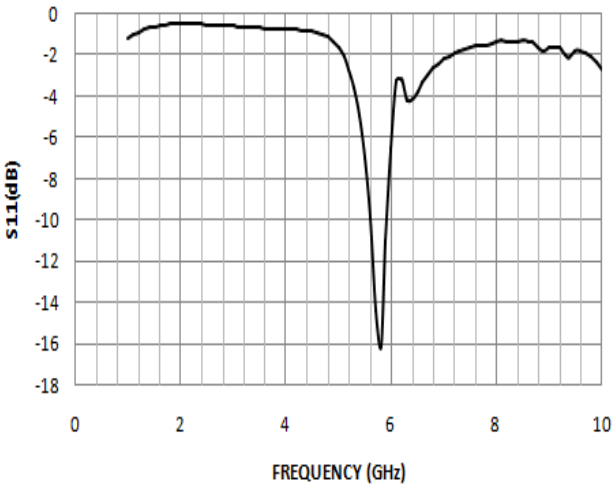


Fig. 6. Simulated S_{11} versus frequency of DRA (dielectric resonator antenna) with cylindrical holes, resonating at 5.8GHz

Radiation Pattern

The radiation plot of the microstrip antenna is normal on the surface of the patch and thus the pattern of elevation is recorded for φ = 0 and φ = 90 degrees. The E-plane and H plane radiation pattern of the plain microstrip antenna at the second WLAN frequency band (i.e. 5.8GHz) is shown in Fig.7 and it can be seen from the figure that the maximum resonant frequency gain is nearly 3.5dBi. Fig. 8 displays the dielectric resonator antenna radiation pattern and it is observed that the maximum 4.7dBi gain is obtained at the resonant frequency. Fig.9 shows the dielectric resonator antenna radiation pattern, which provides an improved gain of 4.79dBi. Therefore the design of two cylindrical holes in a dielectric resonator antenna obtains an improved gain characteristic.

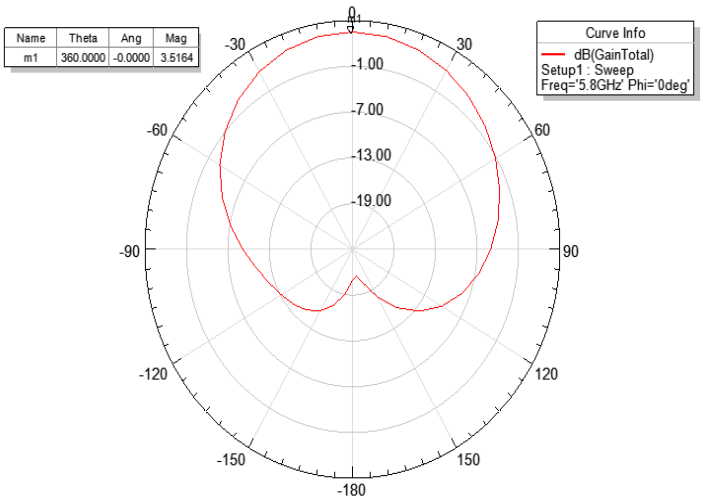


Fig. 7. Radiation plot of the proposed microstrip antenna , resonating at 5.8GHz

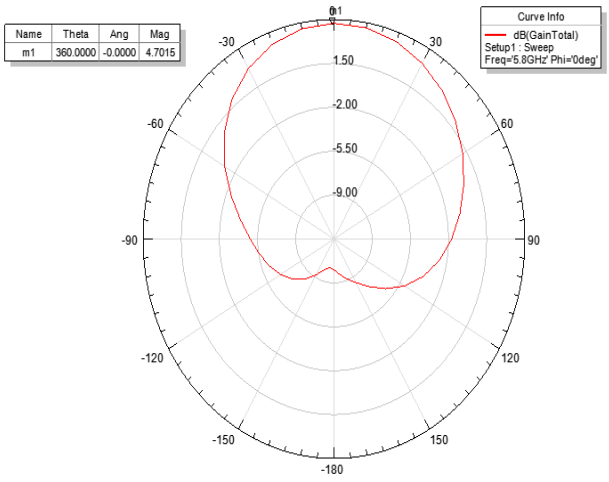


Fig. 8. Radiation plot of DRA (dielectric resonator antenna), resonating at 5.8GHz

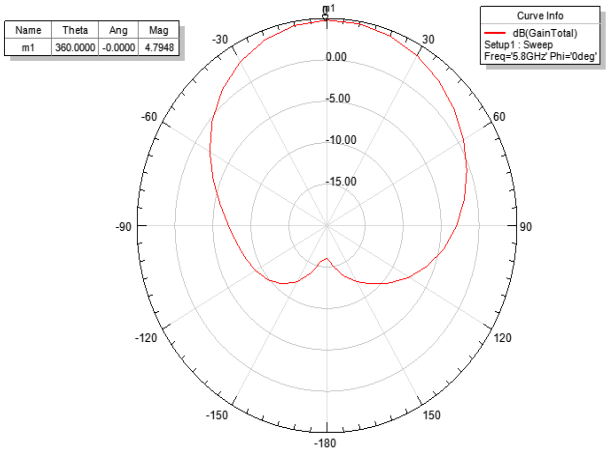


Fig. 9. Radiation plot of DRA (dielectric resonator antenna) with cylindrical holes, resonating at 5.8GHz

Radiation efficiency Versus frequency graph

The radiation efficiency of an antenna can be given as the ratio of the total power that an antenna radiates to the net power received by the connected transmitter antenna. Typically, the radiation efficiency of a microstrip antenna is quite small (< 70 per cent), but improved radiation efficiency is achieved with the implementation of the dielectric resonator antenna in the proposed research. The radiation efficiency versus frequency graphs are shown in Fig.10, Fig.11 and Fig. 12 respectively for the basic rectangular microstrip antenna, dielectric resonator antenna and dielectric resonator antenna with cylindrical holes. From the figures, the radiation efficiency of the simple rectangular microstrip antenna is 60%, the radiation efficiency of the dielectric resonator antenna is approximately 93% and the dielectric resonator antenna is approximately 77%.

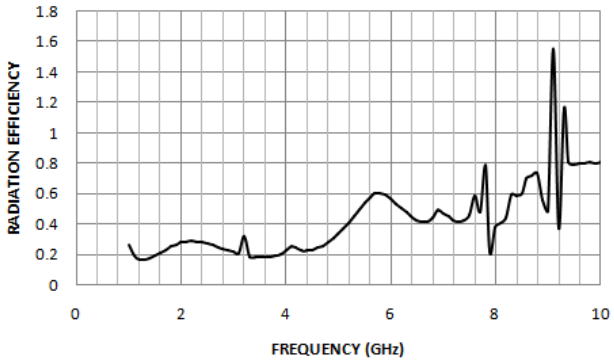


Fig. 10. Radiation efficiency versus frequency of the proposed microstrip antenna, resonating at 5.8GHz

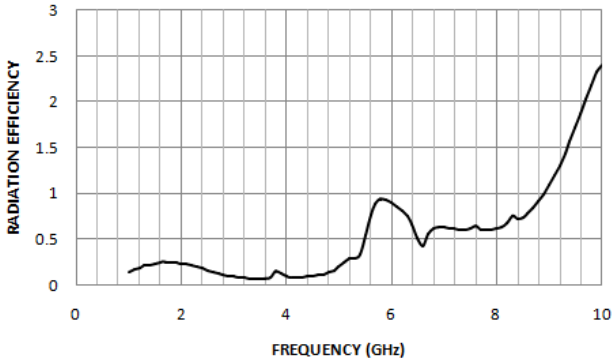


Fig. 11. Radiation efficiency versus frequency of DRA (dielectric resonator antenna), resonating at 5.8GHz

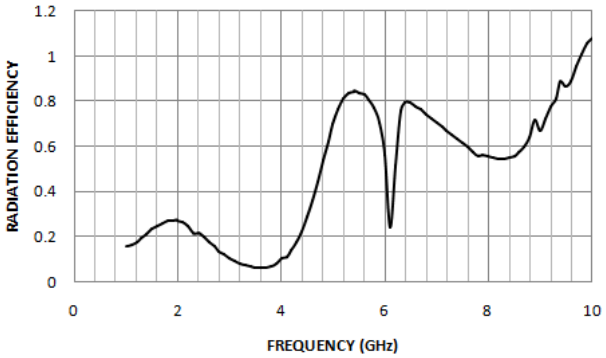


Fig. 12. Radiation efficiency versus frequency of DRA (dielectric resonator antenna) with cylindrical holes , resonating at 5.8GHz

Cross polarization and co polarization

Over time, polarization can be defined as the direction and magnitude variation of the electric field vector.Cross polarization is the orthogonal polarization to the wave 's given polarization, radiated by the antenna, and co polarization is the wave's desired polarization to be radiated by the antenna. For the said plain rectangular microstrip antenna, dielectric resonator antenna, and dielectric resonator antenna the difference between cross polarization and co polarization is shown in Fig.13, Fig. 14 respectively, and Fig.15. From the figures it is observed that the difference in cross polarization and co-polarization is 25dB, 11.5dB and 44dB respectively for the three antennas.

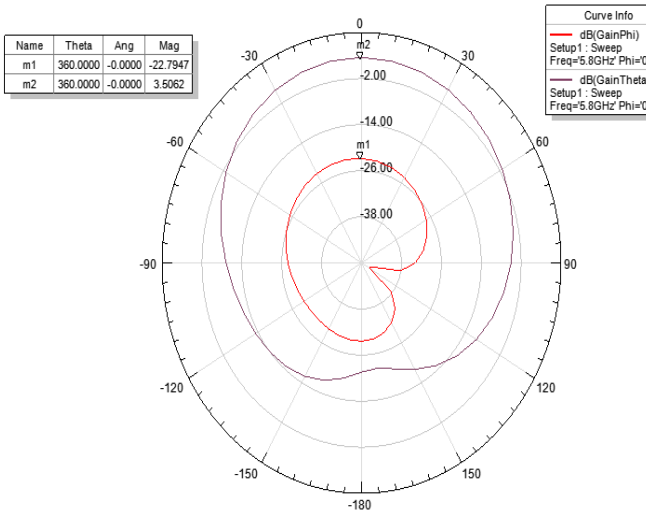


Fig. 13. Cross and co polarization of the proposed antenna , resonating at 5.8GHz

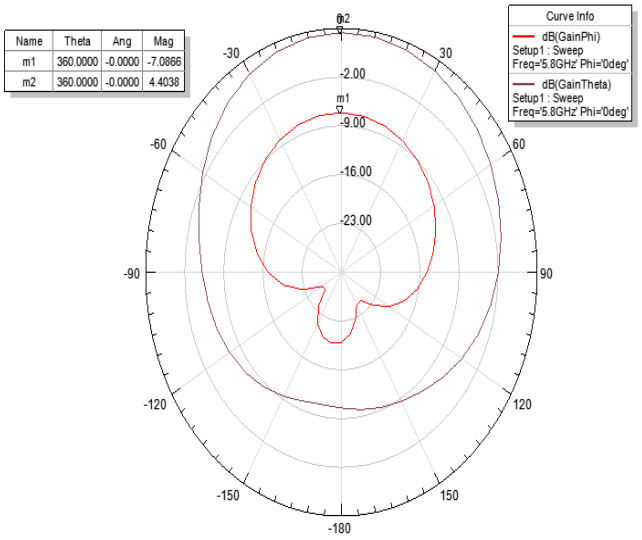


Fig. 14. Cross and co polarization of DRA (dielectric resonator antenna), resonating at 5.8GHz

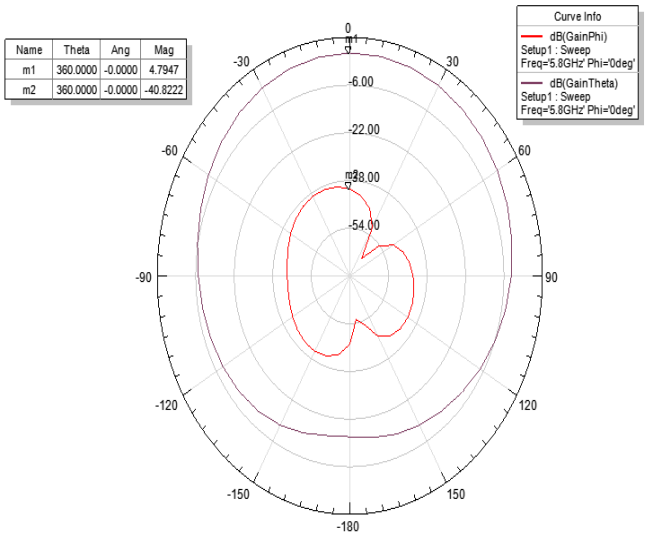


Fig. 15. Cross and co polarization of DRA (dielectric resonator antenna) with cylindrical holes, resonating at 5.8GHz

IV. CONCLUSION

A novel dielectric resonator antenna is proposed, from which gain and efficiency improvements can be achieved. A rectangular microstrip antenna with microstrip feed is designed to be resonant at 5.8GHz, with gain of 3.5dBi and radiation efficiency of 60 per cent. The microstrip antenna then positions a cylindrical structure, which forms a dielectric resonator. Enhanced gain of 4.7dBi and appreciable radiation efficiency of 93 per cent is observed in the proposed dielectric resonator. The concept of a dielectric resonator antenna was introduced to further increase the gain from the dielectric antenna. So two cylindrical holes over the rectangular microstrip antenna were installed in the cylindrical structure, which records an enhanced gain of about 4.79dBi. From the proposed work it can be observed that the gain and efficiency of the microstrip antenna can be significantly enhanced with the use of the dielectric resonator. In this regard, further research is needed, by which the dielectric antenna can be used suitably for practical wireless applications.

ACKNOWLEDGMENT

The authors would sincerely appreciate the encouragement and assistance provided by the faculty members of the Institute of Engineering & Management, Kolkata and Dr. B.C. Roy Engineering College, Durgapur and thus acknowledge the contribution of all concerned.

REFERENCES

[1] K. M. Luk and K. W. Leung ,"Dielectric Resonator Antennas'', RESEARCH STUDIES PRESS LTD. Baldock, Hertfordshire, England

[2] Debatosh Guha, Bidisha Gupta, Chandrakanta Kumar, and Yahia M. M. Antar,"Segmented Hemispherical DRA: New Geometry Characterized and Investigated in Multi-Element Composite Forms for Wideband Antenna Applications", IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, VOL. 60, NO. 3, MARCH 2012

[3] Debatosh Guha, Bidisha Gupta, and Yahia M. M. Antar ,"Hybrid Monopole-DRAs Using Hemispherical/ Conical-Shaped Dielectric Ring Resonators: Improved Ultrawideband Designs'', IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, VOL. 60, NO. 1, JANUARY 2012

[4] Debatosh Guha, Archita Banerjee,Chandrakanta Kumar and Yahia M. M. Antar, ''Higher Order Mode Excitation for High-Gain Broadside Radiation From Cylindrical Dielectric Resonator Antennas'', IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, VOL. 60, NO. 1, JANUARY 2012

[5] Debatosh Guha and Yahia M. M. Antar,'' Four-Element Cylindrical Dielectric Resonator Antenna for Wideband Monopole-Like Radiation'', IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, VOL. 54, NO. 9, SEPTEMBER 2006

A Study on The Financial and Entrepreneurial Risks of Small Business Owners Amidst COVID-19

Andrew Kyung
NVRHS
Demarestl, NJ, USA
kyungandrew314@gmail.com

Sarah Whitney
CRG-NJ
Cresskill, NJ, USA
sarah_whitney@ChoiceRG.com

***Abstract*—The economy suffered a major loss when COVID-19 hit due to business shutdowns. Repercussions from economic downfall due to closed businesses from social distance policies has made many small businesses lose operation mechanisms and see a decrease in profit and sales. Many people lost their jobs due to businesses losses a record 22 million employees have filed for unemployment. In this paper, how the COVID-19 epidemic has changed people's behaviors and changed the way people shop. As more people are shopping online to avoid at risk situations at busy stores, small businesses are seeing losses in sales and customers. Also we studied when they rely on small profits margins, how the small businesses have suffered the worst losses, seeing a decrease in business sales and a lower revenue margin. This research shows how social distance policy has especially affected businesses vulnerable to COVID-19 risks. This often includes businesses that have large groups of people in close quarters or highly susceptible relationships that increase the risks of spreading the illness, such as restaurants or retail stores. Government help has been recommended especially for small businesses and workers who are most vulnerable to economic losses.**

Key Words : stimulus, economy, government, repercussions, epidemic, revenue

I. INTRODUCTION

Most at-risk businesses included those with people who are in close proximity. Second at risk was according to responses to the COVID-19 epidemic. The companies COVID-19 especially affected included: retail trade, transportation, employment services, travel arrangements, arts, entertainment, and recreation, and accommodations and food service. Other businesses could also suffer due to government mandated shut downs. This includes the closure of non-essential businesses that are at risk of spreading the virus. Some of these businesses include museums, movie theatres, gyms, casinos, motion picture and video production, and retailers such as furniture stores, florists, and shopping malls. In many states non-essential also may include construction, certain kinds of manufacturing, and many service sector businesses. About 32 percent or 1.9 million small businesses are at immediate risks of getting closed. Million more are at risk of unemployment.

Many businesses outside the list have chosen to shut down due to risks. Due to the impact of the virus, customer demand has also decreased. Customers feel there is no longer a need for services that are at risk. This includes non-essential businesses such as liquor stores, electronic and appliance stores, day care centers, healthcare facilities and construction. At risk workers include younger, non-white, and non-educated. These are most at risk of unemployment. In order to make businesses more resilient a recovery plan is crucial to the sectors. This may include government funding, new methods of sales such as online, or a change in policy. These businesses must rely on assistance from government programs.

Some companies are especially at risk of the repercussions of the COVID-19 virus, because the illness policy of social distancing has made people avoid busy stores. The social distance policy has greatly reduced demand for the services. Direct contact with businesses has been avoided by customers and businesses have suffered losses of productivity. Many businesses that cannot operate remotely have been forced to shut down due to losses. On average 32 percent of small businesses with fewer than 500 employees or 1.9 million firms have suffered major financial losses. Nearly 20 million employees have been forced out jobs, or led to seek employment in a more secure sector. Mandatory nonessential business shutdowns across the country have a serious impact on small businesses. Vulnerable workers who are already at risk such as younger, non-white, or less educated have been seen to be hit the hardest.

II. FINANCIAL AND ENTREPRENEURIAL RISKS

A. Financial Risks According to Demographic Group

The COVID-19 epidemic has been tough on the economy especially for small businesses and low wage workers. Reports by the Pew Research Center analysis of federal government data show that more than four-in-ten U.S. businesses with paid employees – or 2.4 million out of 5.3 million examined – pose a higher risk in their industry for economic devastation. More then 1 million of these businesses are in retail trade or accommodations or food service because of mandated closing or social distancing policies.

The average American business is small with employees earning about $40,194 yearly. Each business employed an average of 11 workers. Almost three fourths of those businesses or 74 percent earned less than a million a year in sales and about 34 percent had been in business for 5 years or less. Those businesses that remained in higher risks earned less in annual revenue and employed low wage workers that averaged at $28,259 yearly. In economic turmoil, employers

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

and employees of lower income levels have increased risk of facing tough and uncertain economic hardship.

B. Women, Minorities, and the Undereducated

Particular high risk demographic groups such as women, minorities, and the undereducated, often pit in low wage jobs pose a higher threat to economic devastation. On average women are paid less than men, making them more vulnerable to job loss and financial insecurity. According to the culture, White men dominate the job market, owning the majority of businesses. White are over represented in the job market, whereas Black, Hispanic, and Asian pose increased risk of economic insecurity due to lower wage jobs and unstable employment. According to age, younger people lower than 45 years old pose higher risks, often working in lower paying jobs and older people, who are often disposable to employers pose high risks. White men seem to dominate the market and annual revenue of all industries.

Statistics show that lower wage workers are at the highest risk. Industries that employ these workers often include retail and trade, food service workers, manufacturing, entertainment and recreation, and small businesses. Due to mandated shut downs and social distance policies many of these workers had to look for employment elsewhere. This issue with that is that many of these workers are unskilled, undereducated, and underpaid. Finding new employment in a pandemic gets difficult when all businesses are facing annual loss. When pay is low, and there often is not extra money, many people could run out of savings or run low on funds, requesting government aid to lift them out of hard times. Millions of these workers during the pandemic filed for unemployment. This is because they rely on their weekly pay to pay expenses such as food, housing, and transportation. Without government help, these people will not make ends meet.

Particularly unskilled low wage workers are over presented in risks. Millions of these workers faced unemployment and a scarce job market that is not hiring. Small businesses with low annual revenue were also at risk. Business loss in annual profits included job lay-offs and applying for government grants. Business sales plummeted because mandated closing laws in non-essential industries and lower customer rating due to social distance policies. Many people began to shop online. For businesses who could not choose online retail, they saw losses. Businesses had to shut down, customer demand decreased, and annual revenue plummeted. At risks businesses included non-essential products and services, and those with low annual revenue.

The COVID-19 pandemic has also brought racism to various minorities to light. During the spread of the virus, particular inequality towards African Americans (Black Lives Matter Movement), Asian Americans, and other forms of xenophobia were put on display across the media.

The Black Lives Matter Movement was “founded in 2013 in response to the acquittal of Trayvon Martin’s murderer”. In 2020, this movement gained the most global attention in history after the murder of George Floyd. From protests to marches to petitions, various ways of spreading awareness was put across media platforms. Soon enough, a holiday, Juneteenth, has been recognized by various parts of the world as a way to honor black lives.

C. Possible Solutions for Struggling Entrepreneurs

The corona virus has influenced people to stay at home, avoiding shopping at local small businesses. This loss of customer turnout has affected the profit margins of businesses across the country. The small business economy is essential to this country, employing 58.9 million people or about 47.5% of the total private sector workforce. Their GDP average contribution qualified at $5.9 trillion dollars in 2014, when averages were most recently available. Small businesses are productive to the society in that they employ locals, support the local economy, and offer much needed services to the general public. Many local people are employed by small businesses and businesses support the local economy through profits and payroll. The problem is that when there are no customer demands, there’s no money: no money to pay employees, no money to create a supply of customer demanded products, and no relief to slowly recover from economic devastation. This had led many small businesses to be forced to lay off workers, not preparing themselves for recovery when the virus has passed. Because of such economic devastation and the inability to recover, many small businesses will be forced to shut down operations and close for good. This could be catastrophic for the United states economy who relies on small businesses for so many things.

One of the obvious challenges to small businesses is annual revenue, making opening a business risky, and small businesses particularly vulnerable because of the uncertainty of profit. Small businesses have daily costs such as rent, payroll, utilities, and other, making the likelihood of their survival up to chance. Losing revenue from the pandemic could economically devastate them. Due to this, small businesses need to take financial actions in order to stabilize. Under a united states government proposal, the “Small Business Workforce Stabilization Fund,” the treasury would forgive financial assistance that was awarded to small businesses that were secure before the pandemic, as long as the same amount of employees are hired following the crisis. This secures immediate solutions to the most vulnerable businesses, keeps workers getting paid, and inspires businesses to grow when the problem is resolved, by keeping a steady cash flow to all small businesses. This provides security and relief to businesses, employees, and families.

Every business has expenses. Major expenses include paying employees for their services. Large loans can also place pressure on small businesses, especially when the Corona virus crisis hit and businesses were forced to lay off many employees. Having high costs for productivity and

supply, but a loss of demand for the product, caused many businesses to fall into debt. They couldn't pay workers and had to let them go. Businesses needed relief to pay employees, cover expenses, and pay back loans. The $300 billion Restoring Economic Security, Confidence, and User Endurance (RESCUE) Business Act of 2020 would aid in recovery of small businesses. Under this proposed bill, SBA would waive all fees for all 7 loans for one year for both borrowers and lenders, and provide a 90 percent loan guarantee for all loans, no matter the size. Loan limits will be increased from $350,000 to $1 million. This will give businesses time to recover and also keep employees on payroll.

In order to create different policies that work for the benefit of small businesses in Washington, consumers, workers, and job owners need to engage with policy holders. This may include Using social media to spread positive messages, emails, phone calls, and grassroots campaigns that spread the message that small businesses are essential to our lives and a groundwork of the economy. Therefore, engaging to promote relief for small businesses and encouraging government officials to aid, fund, and support the life of small businesses in all communities is essential. Talking to policyholders, writing messages to Congress, and engaging with community members who support the same reforms are essential to creating a robust economy again.

D. Analysis

It is of the utmost importance for non-essential businesses across the country to understand their economic risks laid out before them during these unprecedented times so that they may begin to strategize for the future of their businesses and how to weather this pandemic storm. According to Lexis Nexus (Lexis Nexis, 2020), "economic risk centers on macroeconomic circumstances that may result in significant loss for a business.

The Aggregate Effect has been one that took the most toll on small businesses. Due to the sudden hit of the pandemic, the government has forcefully closed many "non-essential" businesses on short notice. Those who refused risked "citations, fines, or license suspensions" [4]. This forceful closure has taken away some people's only income, especially small non-essential businesses, such as retail or clothing stores.

Though various sources differ in strategies for post-coronavirus comebacks, they all share one factor in common: planning ahead. Specifically, any changes to products or manners of distributing/selling products is of utmost importance, such as assessing or increasing "production and/or delivery capacity" [1]. Whether it's changing the product sold completely, or adjusting advertising and selling techniques, such actions will all be affected by the end of the pandemic.

Being prepared and planning ahead is especially crucial to smaller businesses that don't have the publicity or budget to expand much into online marketing. Even now, organizations such as the Great American Takeout aim to support small businesses, specifically restaurants, during these trying times [10]. By encouraging social media platforms and everyday citizens to order and give publicity to these restaurants, they will not be required to face the burning consequences of funding or other burdens after the pandemic.

I. IV. Discussion

Emergency relief in an economic crisis and immediate aid from government officials and policy makers is necessary to help small businesses survive the pandemic and also recover when necessary. Calls from the general public have encouraged officials to intervene and offer aid in times of crisis. Many programs on State and federal levels have been put into place in order to offer relief to United States citizens. Some of these include grant and loan programs, deferred taxation, a stimulus package, eviction and foreclosure moratorium, utility shut off moratorium, and a call to reduce rent payments. With all these steps policymakers hope to help businesses and employees survive the Corona crisis.

The United States Congress passed a $2 trillion relief package for businesses in crisis from the COVID-19 epidemic. This will be a life saving act that helps millions of businesses recover from the economic crisis of the virus. The future plan includes paying the wages for workers who stay employed at vulnerable companies. For those workers who have lost their jobs, they will receive checks as large or even larger than they were receiving at previous employment. It will save some of the country's largest corporations from bankruptcy, by allowing taxpayers to take shares and invest in those companies as collateral. This is the first time the government has actively involved itself in capitalistic business. This bill will transfer responsibility from private businesses to the federal government, allowing the government to intercede on behalf of companies and potentially aid the recovery of the pandemic.

The effects of the corona virus have sent millions of Americans into unemployment. Millions of companies both big and small are struggling to survive the financial wreckage of the virus, including airlines, big banks, nail salons, and bars. This has led the United States government to intervene on behalf of businesses. Some of these provisions include direct payments to individuals and businesses, generous loans in which the government agrees to support equity and stakes in companies. Some individuals will receive a check to compensate for lost wages. But, there's provisions such as a limit on executive pay and a requirement of companies receiving assistants to maintain employment rates of 90 percent of what they were.

The government has planned to help many businesses. Around $350 million will be devoted to loans and loan payments. The other $500 billion will be divided among airlines and other companies critical to national security. The government has done many things to aid in the recovery.

There are some loopholes to their cash bailout. Businesses would not have to repay loans covering up to eight weeks of payroll expenses. Once businesses receive their loans, a new payment schedule would begin. They must use the money within two months in order to avoid repaying it. They also cannot pay employees more than $10 thousand if that is the size of their loans. Executive who made $3 million, will receive $3 million. CEO's who made $20 million would receive $11.5 million. Restrictions will be supplied when government support begins.

Direct payments, low interest loans, bailouts, and deferred repayments, all will aid in post-pandemic recovery for businesses. Businesses who have lost revenue, lost employees, and struggle to recover, will now get the help they need to survive the economic crisis. When the crisis hit, the government made it their responsibility to provide much needed aid to companies that have found that they have lost profits, income, and production. Many companies folded, employees were let off, and daily profit margins fell substantially. The pandemic causes much stress and economic hardship for many people. Government intervention was necessary to recover and keep businesses working. As the government has chosen to intervene, and aid in the recovery of companies through economic hardship.

The Corona Crisis has left many people desperate for cash. As people are losing their income rapidly, there was an urgent need for the government to intervene and give them a money boost. At the public's urgency, the government issued its first stimulus package which is a refundable payment against your 2020 tax returns. This means that you don't have to file your 2020 tax returns first to get the tax credit in the form of a tax refund. The stimulus payment will not be taxed by the government. At a desperate time, this is just what many people needed. With businesses at a loss and millions of people becoming unemployed, people have been unable to pay rent and mortgages, buy food, or have transportation. In order to pay expenses,

This was the first of a number of stimulus checks being issued and just a part of the government's plan to aid in the economic recovery of the American people. When the economic crisis hit, a plan was necessary to give people relief from economic hardship. In order to help people, the government issued checks of $1,200 to give people aid to pay rent, mortgages, and buy food. Many people have been speaking about ending rent payments in order to give people relief. They also talk about ending evictions. Others spoke about a second stimulus check in order to help because the corona virus went on for so long. Direct money is what people need to pay expenses. A stimulus check was the government's solution to the problem of economic hardship.

People have demanded that the government intervene even more. People who are low income and already in crisis can not afford unemployment. The average worker who makes less than $20,000 a year and lives on paycheck to paycheck cannot afford to lose any more income. Unemployment, loss of income, and economic hardship will devastate them because their savings are small and the money is dwindling. In many ways, these people have required the government to intervene in their livelihoods. One crisis can set these people off to homelessness and leave them empty. Many people have gone to unemployment benefits and other aid, but found with high costs, it's not enough. A stimulus check is necessary for intervention and for people to pay necessary expenses like food, shelter, and transportation.

REFERENCES

[1] Monson, C. (2020). What small Businesses Need to Survive The Corona Crisis. Harvard Business Review.

[2] Ding, L., Sanchez, A. (2020). What small Businesses Will Be impacted by COVID-19? Federal reserve Bank of Philadelphia.

[3] Kochhar, R. (2020). The Financial Risk to U.S. business owners posed by COVD-19 outbreak varies by demographic group. Pew Research Center.

[4] Smith, K. (2020). How State and Local Governments are Helping Small Businesses During The Coronavirus Pandemic. Institute For Local Self-Reliance.

[5] Tankersley, J (2020). The Corona Virus Economy: When Washington Takes Over Business. The New York Times.

[6] Davidson, Helen (13 March 2020). "First Covid-19 case happened in November, China government records show—report". *The Guardian*. ISSN 0261-3077

[7] Sevillano EG, Linde P, Vizoso S (23 March 2020). "640,000 rapid coronavirus tests arrive in Spain". *EL PAÍS*.

[8] Meredith S (20 March 2020). "UK PM Boris Johnson announces nationwide lockdown measures, telling cafes, pubs and restaurants to close". CNBC.

[9] "Principles of Epidemiology | Lesson 3—Section 3". *U.S. Centers for Disease Control and Prevention (CDC)*.

[10] Gehanno JF, Bonneterre V, Andujar P, Pairon JC, Paris C, Petit A, et al. (June 2020). "How should data on airborne transmission of SARS-CoV-2 change occupational health guidelines?". *Occupational and Environmental Medicine*.

[11] Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. (April 2020). "The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak". *Science*.

Telemetry System for Experimental Study of Ultrasonic Measurement on COVID-19 Situation

Wongsakorn Wongsaroj
Department of Mechanical Engineering
Tokyo Institute of Technology
Tokyo, Japan
wongsaroj.w.aa@m.titech.ac.jp

Naruki Shoji
Department of Mechanical Engineering
Tokyo Institute of Technology
Tokyo, Japan
shoji.n.aa@m.titech.ac.jp

Hideharu Takahashi
Laboratory for Advanced Nuclear Energy, Institute of Innovative Research
Tokyo Institute of Technology
Tokyo, Japan
htakahashi@lane.iir.titech.ac.jp

Hiroshige Kikura
Laboratory for Advanced Nuclear Energy, Institute of Innovative Research
Tokyo Institute of Technology
Tokyo, Japan
kikura@lane.iir.titech.ac.jp

***Abstract*—As the COVID-19 situation, the social gathering and working activity is limited. The research activities and engineering work has been affected strongly especially the experimental work. In this paper, the telemetry system concept is proposed for conducting the experiment remotely outside the laboratory. The power source of the experimental equipment is controlled remotely. The experimental execution is observed by the camera exposing the user to monitor the real time conducting from the outside. The experiment of velocity profile measurement on the bubbly flow which is the importance task in the fluid engineering was demonstrated to confirm the applicability of the concept. Finally, the experimental result can be obtained successfully under the remoted experiment.**

Keywords—Telemetry, COVID-19, Experiment, Ultrasonic,

I. INTRODUCTION

Nowadays, among the high demand of the electrical energy consumption in the world. The power plant is established to respond this requirement. The one of electrical power generation is based on the utilizing of steam power. This concept is placed on coal, gas combine cycle and nuclear power plant. However, these power plants must be operated safely without of the accident. The operation is imperative to under the safety aspect. The crucial part of the plant operation according to safety standard and sustainability is the engineering design. The thermal hydraulic and fluid mechanic behaviors of the water-steam cycle in the boiling unit of steam generation is the main section on the occurring of the accident. It is compulsory to be predicted and optimized accurately by the numerical model [1, 2]. These behaviors directly determine the geometry, material and other specification which affects the safety criteria of the unit. To confirm reliability of the model, Indispensably, the model validation by the experimental data is the important task. Hence, experimental investigation on this behavior is needed [3].

Presently, due to the COVID-19 situation, the social gathering and working activity in the work place have been prohibited or at least is minimized. Significantly, it influences to the activity on the experimental task in the facility placed on the laboratory or company. As the consequence, the completion of experiment work is delay. Then, the engineering design and project execution as well is belated. Hence, to solve this problem, the conducting of experimental work remotely from other places outside the work place has been considered.

The internet of thing (IoT) [4] is a system of interrelated computing devices, mechanical and digital machines provided with unique identifiers (UIDs). It has the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction. As this technology, it opens the opportunity for remoted operation, automation and telemetry system in several sectors such as industrial [5], medication [6] or even in for smart home [7]. Therefore, this platform has a potential significantly to apply for the execution of experimental work remotely under pandemic disease circumstance as the COVID-19.

This study presents the telemetry concept to hand on the experimental activity remotely. The experiment can be conducted anywhere at the internet is available. The demonstration is focused on the experiment of the velocity distribution measurement in the bubbly flow. It is the one of important parameter of fluid mechanic investigation. The investigation is conducted by employing the ultrasonic velocity profiler measurement system.

II. MEASUREMENT DEVCE AND TELEMETRY SYSTEM

A. Ultrasonic velocity profiler

The UVP measurement [8] is an ultrasound-reflection technique which can obtain the velocity profile of the liquid. Figure 1 illustrates the UVP configuration, ultrasonic pulse-echo and velocity profile reconstruction. An ultrasonic pulse is transmitted from the transducer along the measurement line to the liquid. The echo signal reflected from moving reflector such as small particle dispersed in the liquid is derived by the same transducer. As the ultrasonic wave is emitted repeatedly, the echo signals are obtained sequentially. Doppler signal influenced by the velocity of a moving particle can be de-modulated from the echo signals. The Doppler frequency $f_D(i)$ directly relates to the velocity of the particle (i) is position or channel). Hence, the velocity of the particle at that position $V(i)$ can be computed as

$$V(i) = \frac{cf_D(i)}{2f_0 \sin\theta} \quad (1)$$

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

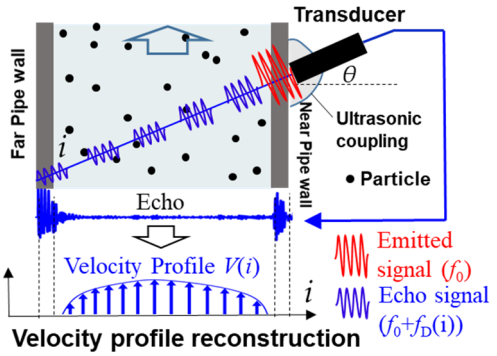


Fig. 1. The UVP measurement configuration.

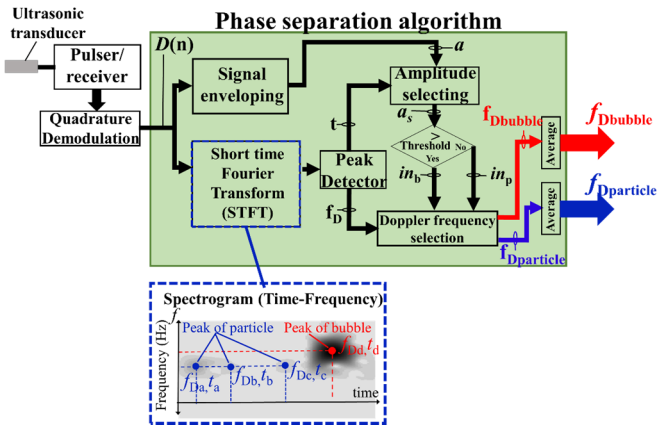


Fig. 2. The UVP with phase separation for measurement in the bubbly flow.

where f_0 is the basic frequency and θ is the incident angle. If the stokes number on the relation between small particle and the liquid < 0.1, the particle will follow liquid streamline closely. Then, several particles dispersed in the liquid. Consequently, the velocity profile of the liquid can be obtained.

In the bubbly flow, the phase separation technique for the operation in the UVP system to obtain the velocity in the bubbly flow was proposed [9]. Figure 2 illustrates the process of proposed algorithm. Firstly, the Doppler signal $D(n)$ (discrete data) is extracted from the echo signals $e(t)$ as shown in Equation (2) obtained from transducer and pulser/receiver respectively. The extraction is processed by quadrature demodulation [10].

$$D_i(n) = A_{n,i}\cos\left(\frac{2\pi n f_D}{f_{PRF}} - \varphi\right) - iA_{n,i}\sin\left(\frac{2\pi n f_D}{f_{PRF}} - \varphi\right) \quad (2)$$

where n represents sampling of Doppler signal, A is the amplitude, and φ is the initial phase. Then, the Doppler signal is sent to STFT to derive time-frequency spectrogram of signal. The calculation is expressed in Equation (4) and the energy density of spectra at time k is denoted by Equation (5). Time-frequency resolution depends on time step S_n and window length W_n. The spectrogram is sent to the peak detector for analyzing the energy peaks of the spectrogram. Peak value in each position informs the Doppler frequency data ($\mathbf{f_D}$=[f_{Da},f_{Db}...f_{Dm}]) and time location ($\mathbf{t}$=[t_a,t_b….t_m]). Furthermore, the Doppler amplitude at each point ($\boldsymbol{a}$=[$a_{n=0}$, $a_{n=1}$, ..$a_{n=Nrep-1}$]) is detected by making the envelope on the Doppler signal. These data are selected by time location index obtained from peak detector ($\boldsymbol{a_S}$=[a_a, a_b,...a_m]).

$$X(k, f_D) = \sum_{n=0}^{N_{REP}-1} D(n)W(n-kS)\exp(-jn2\pi f_D) \quad (3)$$

$$P(k, f_D) = \left|X(k, f_D)\right|^2 \quad (4)$$

The selected amplitude data is then compared with a threshold value. The value is defined as being higher than the maximum Doppler amplitude of the particle and lower than the Doppler amplitude obtained from the bubble. The amplitude index is classified into the index of bubble and liquid. When the amplitude value is higher than the threshold, the index is defined as a bubble index ($\mathbf{in_b}$=[in_{b1},in_{b2}..in_{bn}]). Furthermore, when the value is lower than the threshold, the index is expressed as a particle index ($\mathbf{in_p}$=[in_{p1},in_{p2}..in_{pn}]). Doppler frequency data analyzed by peak detector is classified by these amplitude indexes to be Doppler frequency of bubble group ($\mathbf{f_{Dbubble}}$ =[f_{Dba},f_{Dbb}…,f_{Dbm}]) and particle group ($\mathbf{f_{Dparticle}}$ =[f_{Dpa},f_{Dpb}….f_{Dpm}]). The Doppler frequency in each group is averaged. Hence, the Doppler frequency of bubble $f_{Dbubble}$ and particle $f_{Dparticle}$ in the same measurement channel is decomposed apparently. Consequently, the velocity of the bubble and particle (liquid) can be calculated simultaneously follow the equation (1).

B. Integrated telemetry system

The internet of things (IoT) is a hot issue and plays an important role in several sectors [11]. For the telemetry system applied for the experimental work remotely, the client units that work and facilitate the remote work execution communicates with the outside users via Wi-Fi signal internally and with internet services provider (ISP) respectively as shown in figure 3. The ISP provides internet service that connect data sent through router devices such as Dial, DSL and wireless modems where the latter is selected for this study. The computer, tablet and smartphones are responsible for remote control over the internet anywhere via 3G/4G or routers that communicate with the ISP.

III. EXPERIMENTAL SETUP

Figure 4 illustrates a schematic diagram of the measurement system and the experimental apparatus. The UVP system consisted of 4 MHz transducers, an ultrasonic pulser/receiver, a digitizer, and a computer with LabVIEW software. The pulser/receiver emitted ultrasonic pulses via a transducer. The echo signals received by the pulser/receiver were converted into a digital signal by the digitizer, with a sampling rate of 100 MS/s. Data from the digitizer were transferred to the computer via USB port. The UVP calculation and analysis were performed by using the LabVIEW software on the experimental computer. The UVP parameter setting shown in table 1 was set to be proper with the measurement condition. For the demonstration, the UVP was applied to measure the velocity profile of the single phase liquid and air bubbly flow on the vertical pipe flow apparatus. The tap water dispersed with nylon particles 80 µm and bubbles (diameter ≈ 2-3 mm) were used as working fluid. Its temperature is 25± 4ºC. The water was circulated by the pump. The water flow rate was monitored by the electromagnetic flow meter. Bubbles were generated by a bubble generator

TABLE I. PARAMETER CONFIGURATION OF UVP SYSTEMS

Parameter	Value
Basic frequency	4 MHz
Pulse repetition frequency	8 kHz
Number of repetition	128
Channel width	0.74 mm
Sound velocity in water	1493 at 25 degree C
Incident angle	45 degree

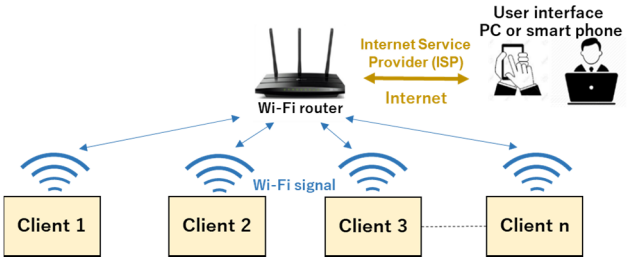


Fig. 3. Schimatic of communication network for remoted work operation

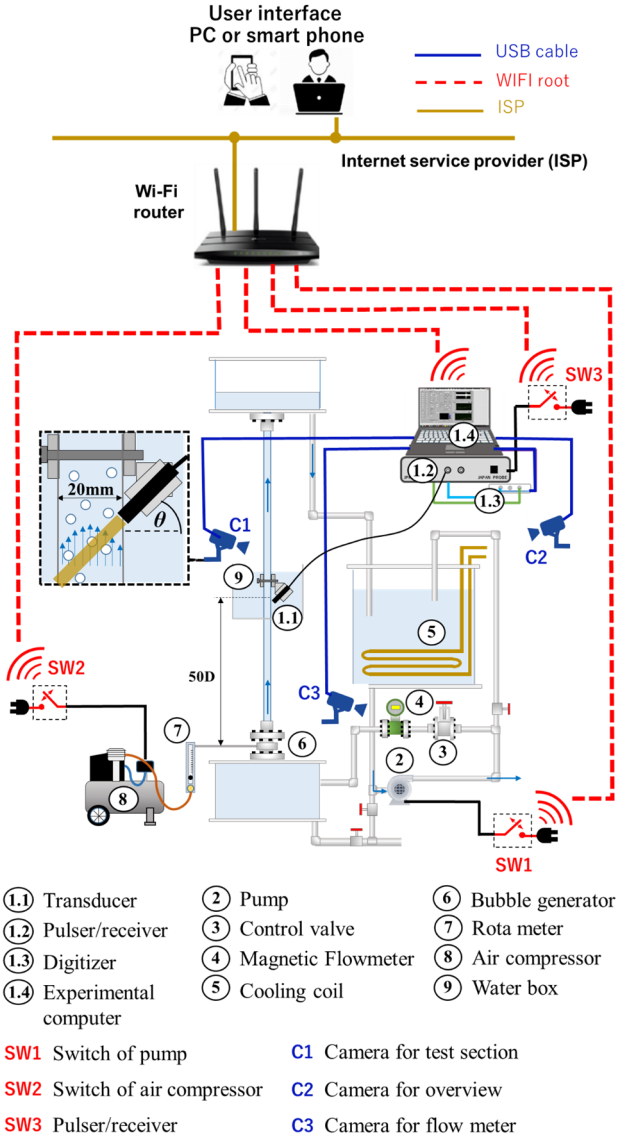


Fig. 4. Experimental setup.

put on upstream 50D from test section. The transducers were immersed into the water box. The transducer end was put on the outside of the pipe with the incident angle 45 degree.

In order to control the experimental work remotely, the web camera, smart electrical switch and the remote software were employed in the experiment. Three cameras and two smart switches were worked on the system. First web camera was used for overview observation. Second one was used to monitor the flow rate value on the flow meter. other one was installed to observe on test section area. The smart switches were utilized to control the power supply of pulser/receiver, water pump, and air compressor separately. The video data transferring of web camera was connected to the experimental computer via USB port and its data was transferred by via Wi-Fi respectively. The switch power control were executed via Wi-Fi network. The information was communicated with user interface by utilizing the ISP. This experiment system were proposed to operate via the remote access tool (Team viewer software was employed).

IV. RESULT AND DISSCUSSION

A. *Telemetry system testing*

Figure 5 shows the web browser of smart switch control. These switches was used to control the power source to energize the pulser/receiver, fluid circulating pump, and air compressor of bubble generator. The power of equipment can be controlled remotely on the web browser. Besides, the remoted observation during the experiment was operated via web cameras as illustrated in the figure 6. The flow meter display unit was monitored continuously for accurate experiment. The test section and overall view of the apparatus were observed due to the safety aspect. Figure 7 shows the operating panel of the UVP program. The program was developed on the LabVIEW version 2011.

Figure 8 illustrates the remote access screen of the telemetry system. Every browser was arranged to fit the monitor.

B. *Experiment in liquid flow measurement*

UVP parameters such as basic frequency, pulse repetition frequency, measurement channel, number of profile and others could be set via the remote access software. Figure 9 represents the results of averaging data of 5,000 instantaneous velocity profile in single-phase liquid flow at superficial liquid velocity U_L of 300 mm/s. The horizontal axis is the distance

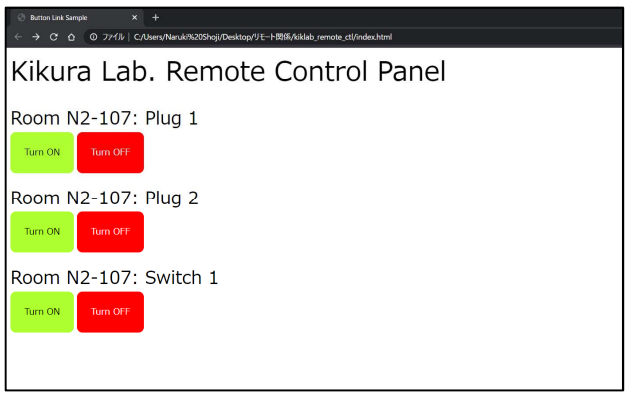


Fig. 5. Smart switch control browser.

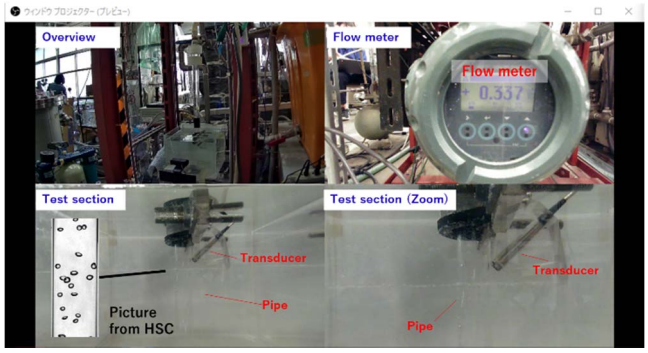


Fig. 6. Camera viewing browser.

from the wall (r) nominalized by the pipe radius (R). The UVP system and liquid pump were started remotely. However, the velocity value of UVP near the vicinity of the wall showed small fluctuations due to some parts of the ultrasonic measurement volume is located within the wall which is the overlapping of the ultrasonic wave and pipe wall. The echo signal in that zone is not only affected by the particle. The measurement result is influenced by moving particles and the pipe wall. Hence, it can be concluded that the UVP efficiently measured the velocity profile on a single-phase flow even executed via the telemetry system. The measurement result can be transferred to the user interface PC promptly for post processing.

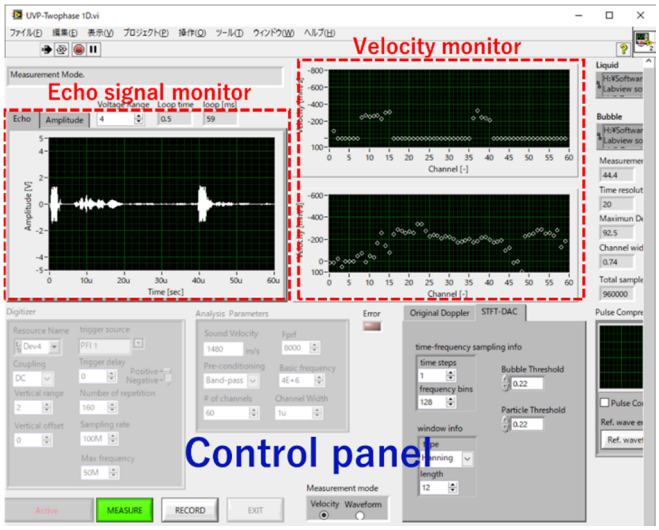


Fig. 7. The UVP program operation panel

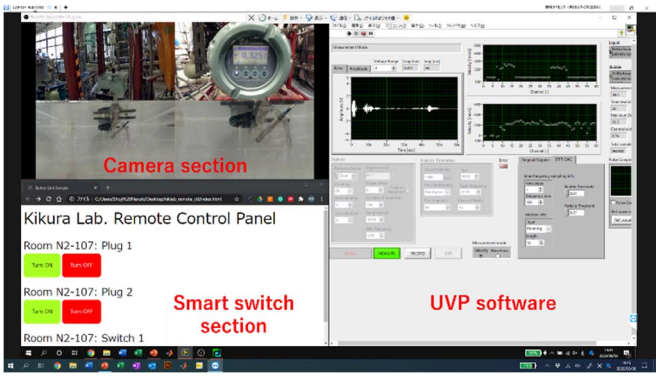


Fig. 8. The screen of the remote access to experimental computer.

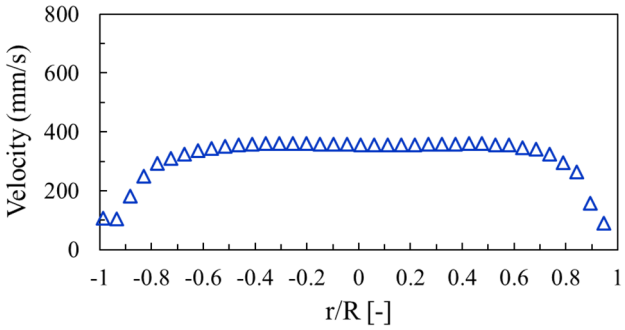


Fig. 9. The result of liquid flow.

C. *Experiment in bubbly flow measurement*

In the bubbly flow, the experiment was conducted at U_L = 300 mm/s and U_G = 5.3 mm/s. The bubble diameter in this experiment was about 2-3 mm. The air compressor was powered up by telemetry system to inject bubbles. Figure 10 represents that the instantaneous velocity profile of the bubble and liquid can be obtained remotely. Figure 11 shows the average measurement results of two-phase bubbly flows. Bubbles rise mainly near the wall region. The graph shows the mean of velocity profile data. Liquid velocity distribution is the averaging of 5,000 profiles and bubble velocity is averaged by the amount of data obtained. The measurement result of bubble velocity profile after separation by the developed technique was derived. The velocity level is higher than liquid velocity due to the buoyancy force affect. Besides, the liquid velocity profile was obtained and separated from the bubble phase obviously. The velocity of both phase can be measured separately. Then, the slip ratio also was derived. It can be summarized that the bubble velocity, liquid velocity and slip ratio of the bubbly flow were investigated although the experiment conducted on telemetry system.

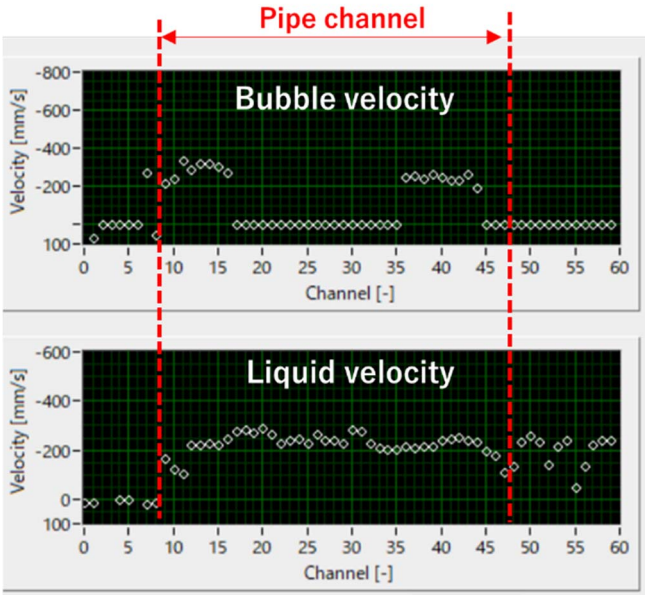


Fig. 10. The result of instantanous velocity on bubbly flow.

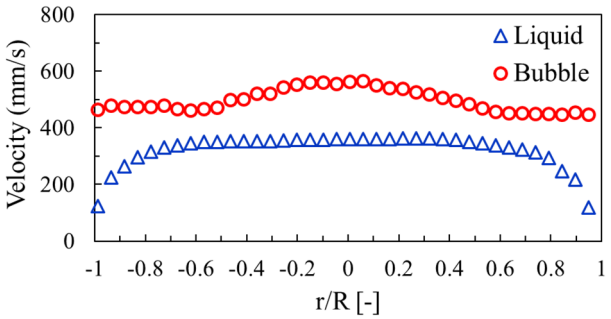


Fig. 11. The result of average velocity on bubbly flow.

V. Conclusion

The telemetry concept to hand on the experimental activity remotely was proposed. The experiment can be conducted anywhere at the internet is provided. Therefore, obstacle of experimental task in crisis or pandemic disease situation is removed. The experiment of the velocity distribution measurement in the bubbly flow using ultrasonic measurement, which is important in fluid engineering, was demonstrated to confirm the ability of the telemetry system. The UVP system, liquid pump and bubble generator were controlled via web browser. The UVP program operation and web cameras were executed with the remote access tool. The results of velocity profile on the remoted experiment whether single phase liquid and bubbly were obtained clearly.

References

[1] M. Ishii, and T. Hibiki, Thermo-fluid dynamics of two-phase flow, Springer, 2012.

[2] T. Ozaki, R. Suzuki, H. Mashiko, and T. Hibiki, Development of drift-flux model based on 8×8 BWR rod bundle geometry experiments under prototypic temperature and pressure conditions, Journal of Nuclear Science and Technology, 2013, vol. 50, pp. 563-580.

[3] S. Hosokawa, K. Hayashi, and A. Tomiyama, Void distribution and bubble motion in bubbly flows in a 4×4 rod bundle. Part I: Experiments, Journal of Nuclear Science and Technology, 2014, vol. 51, pp. 220-230.

[4] L. Atzori, A. Iera, and G. Morabito, The Internet of Things: A survey, Computer Networks, 2010, vol. 54, pp. 2787-2805.

[5] H. Boyes, B. Hallaq, J. Cunningham, and T. Watson, The industrial internet of things (IIoT): An analysis framework, Computers in Industry, 2018, vol. 101, pp. 1-12.

[6] S. Chaudhury, R. Mukherjee, D. Paul, and S. Haldar, Internet of Thing based healthcare monitoring system, 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON), 16-18 Aug 2017, Bangkok, Thailand.

[7] S. Dey, A. Roy, and S. Das, Home automation using internet of thing, IEEE 7th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 20-22 Oct 2016, New York, NY, USA.

[8] Y. Takeda, Velocity profile measurement by ultrasonic Doppler shift method, International Journal of Heat Fluid Flow, 1986, vol. 7, pp. 313–318.

[9] W. Wongsaroj, A. Hamdani, N. Thong-un, H. Takahashi, and H. Kikura, Extended Short-Time Fourier Transform for Ultrasonic Velocity Profiler on Two-Phase Bubbly Flow Using a Single Resonant Frequency, Computers in Industry, 2019, vol. 9, 50.

[10] Y. Takeda, and et.al., Ultrasonic Doppler Velocity Profiler for Fluid Flow: Springer: Tokyo, Japan, 2012.

[11] P. Park, S. C. Ergen, C. Fischione, and C. Lu, Wireless Network Design for Control Systems: A Survey, IEEE Communications Surveys & Tutorials., 2018, vol. 20, pp. 978-1013.

An Application of IoT for Conduct of Laboratory Experiment from Home

Tri Vien Tran
Department of Mechanical Engineering
Tokyo Institute of Technology
Tokyo, Japan
tran.t.ae@m.titech.ac.jp

Hideharu Takahashi
Laboratory for Advanced Nuclear Energy
Tokyo Institute of Technology
Tokyo, Japan
htakahashi@lane.iir.titech.ac.jp

Tadashi Narabayashi
Laboratory for Advanced Nuclear Energy
Tokyo Institute of Technology
Tokyo, Japan
tnaraba@lane.iir.titech.ac.jp

Hiroshige Kikura
Laboratory for Advanced Nuclear Energy
Tokyo Institute of Technology
Tokyo, Japan
kikura@lane.iir.titech.ac.jp

***Abstract*—The concept "Internet of Things" has been applied in many fields to reduce human efforts by using automatic things. It also provides opportunities to develop remote laboratories or distance researches/studies. The main objective of this article is to conduct the telemetry experiment on the measurement of the flow behavior in the pool scrubbing of a Filtered Containment Venting System (FCVS). The Internet of Things (IoT) has been applied to monitor and control the experimental performance and the measurement system. An IoT network that consists of various sensors, web-cameras, IoT appliances, and PCs was established. Thus, the laboratory's experimental works can be implemented from home. It is valuable for the current situation of social distancing due to the Covid-19 pandemic.**

Keywords—Internet of Things (IoT), Filtered Containment Venting System (FCVS), Remote Laboratory, Telemetry research.

I. INTRODUCTION

One of the most important lessons learned from the accident of the Fukushima Daiichi Nuclear Power Plant is that a system should be required for accident management. A Filtered Containment Venting System (FCVS) is one such system that can enhance the capability to suppress or even prevent the occurrence of a severe accident due to the reduction of the pressure, steam water, and flammable gas in the reactor containment vessel. With the equipment of the high-efficiency filtration system, it is also important to minimize the release of fission products and radioactive material to the environment in case of an accident occur [1, 3]. The FCVS is also effective to remove the radioactive aerosols involved during the decommissioning of the reactor due to dismantling facilities and cutting the core debris [2]. After the Fukushima accident, the FCVS has been developed in several countries. The FCVS also become the regulatory requirements for the new constructions of the nuclear power plant [3]. Therefore, understanding the performance as well as improving the efficiency of FCVS is important; and the experimental researches on FCVS has been getting more attention recently.

In the current situation of the Covid-19 pandemic, almost countries have been applied to the social distancing measure to reduce the rapid spread of the virus infection during the vaccine that has not been successfully developed. Working-from-home is a temporary solution for most organizations and universities. However, this solution is not always fit for everyone, especially the students who need to conduct experimental researches. The limitation of the onsite working time may slow down the progress of the study.

With the development of computer technology and communication technology in recent years, the Internet of Things (IoT) concept has been considered as the objective of many research and applications. IoT is a defined network of mechanical, digital, and computing devices that able to collect and exchange data. A typical example of IoT is its application in a smart home such as the security system which consists of a variety of smart appliances and electronic devices that can be controlled remotely via smartphone or a computer connecting to the internet. Besides, the IoT has been also applied in many areas in the industry, agriculture, and healthcare. Together with the development of IoT technology, there are many smart appliances and software has been also developed to allow remote control and monitoring. Therefore, the IoT provides opportunities for remote conducting experimental researches in this situation of Covid-19.

In addition, the concept "Remote Laboratory" has been introduced in recent years as one of the solutions that improve collaborative researches and education. The development of remote laboratories is valuable for distance education and research works. More students/researchers can join in one experiment without staying on-site. It will also help to reduce the study cost in comparison to the traditional laboratories due to the sharing lab and apparatus [4, 5]. The IoT plays a very important role in the operation of such remote laboratories. It supports a communication platform between researchers and experimental devices.

Several works have been done in the application of IoT to perform experiments remotely. Most of them focus on controlling a simple training experiment. For example, controlling a Separately Excited DC Motor [6], testing of proximity sensors [7], or controlling the lab electrical devices [8], etc. However, the engineering fields are varied, and sometimes the experiments require specialized measuring equipment. Unfortunately, most of these devices have no option for remote control mode. The remote control of these devices can be possible by adding some proper smart IoT things.

The major objective of this paper is to measure the two-phase flow behavior in the scrubbing pool of a laboratory-

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

scale FCVS using the visualization method. The IoT network which consists of several web-cameras and smart things has been established for distance monitoring and controlling the experiment as well as collecting the measurement data. It is also the preparations for the telemetry experiments for our projects of decommissioning the Fukushima Daiichi Nuclear Power Plant and establishing the remote laboratory in the near future.

II. Experimental Setup

A. Facility

For the experimental study on the performance of an FCVS, a test facility was constructed in the Tokyo Institute of Technology. The schematic of the FCVS laboratory-scale test facility is shown in Fig. 1. It simulated a multi-layer filtration system including a scrubbing pool wet-type filter, a mist eliminator, and a multi-stages dry-type filter. The multi-stage dry-type filter allows to combine several specific filters for different types of simulated radiation aerosol and gas, such as Metal Fiber Filter for aerosol removal or the Zeolite Filter for removal of organic iodine; etc. In the scope of this paper, only the flow behavior of the pool scrubbing of the wet-type filter was investigated using the visualization measurement method.

The wet-type filter is the first filtration stage of the FCVS. It is a vertical cylindrical scrubbing pool made of transparent acrylic with an inner diameter of 20 cm, the wall thickness of 1 cm, and a total length of 200 cm. The working liquid was used in the wet-type filter is water. A Venturi Scrubber Nozzle was used to generate the air bubble in the pool.

The simulated air venting flow was supplied by the air compressor (ETG Japan Vacmaster VO1220SFD-SP). The venting flow rate can be adjusted by the control valves on the downstream of the air compressor. The pressure and the flow rate of the airflow at the inlet were measured using the pressure gauge and the flowmeter which was installed on the pipeline downstream of the control valves.

The square cross-section Venturi Scrubber Nozzle was used for bubble generation. It was designed as the double-stages Venturi Scrubber with the throat size of 14 mm x 58 mm; and the maximum cross-section (the exit) is 58 mm x 58 mm. At the exit of the nozzle, a baffle plate was installed to break up the large size bubbles and changing the moving direction of bubbles from vertical to lateral when it escaped from the nozzle.

The Mist Eliminator Vane-Type was installed at the downstream of the scrubbing pool. It is also played as the secondary filtration layer of FCVS to remove the water entrainment droplets and aerosol particles that escaped from the first layer without absorbed into the water pool.

B. Measurement System

The flow behavior of the pool scrubbing in the wet-type filter layer was visualized using a high-speed camera (HSC) (Fastcam Mini AX50 manufactured by Photron Co. Ltd.) with the Nikon 60mm f/2.8G lens. The backlight illumination was applied for visualization. A halogen lamp was set up on the opposite side of the pool from the HSC. To reduce the refraction effect which is caused by the cylindrical-shaped of the scrubbing pool wall, a rectangular water box was installed covering the scrubbing pool at the test section as shown in Fig. 2. Because of the similarity of the refractive indexes between

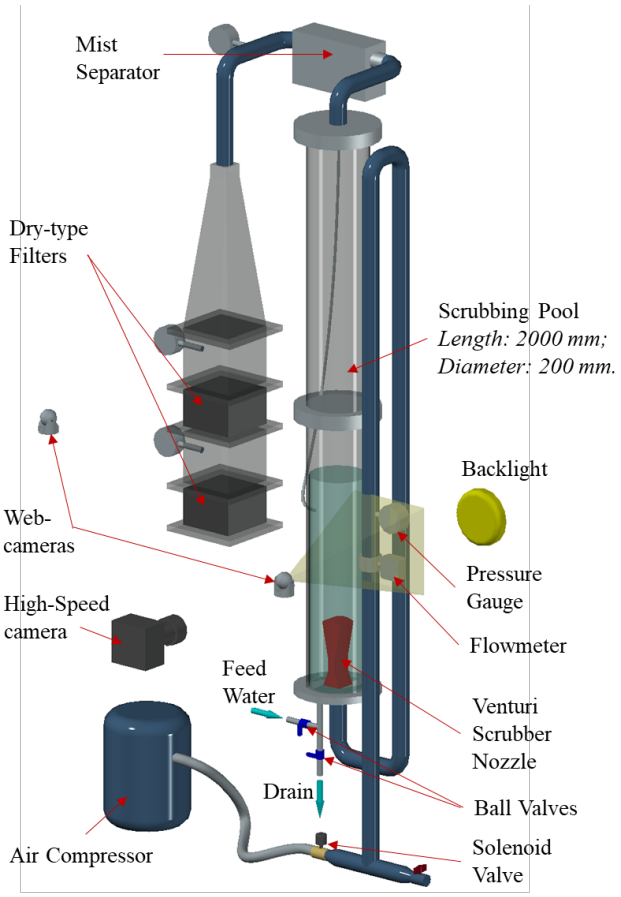


Fig. 1. Schematic diagram of the test facility.

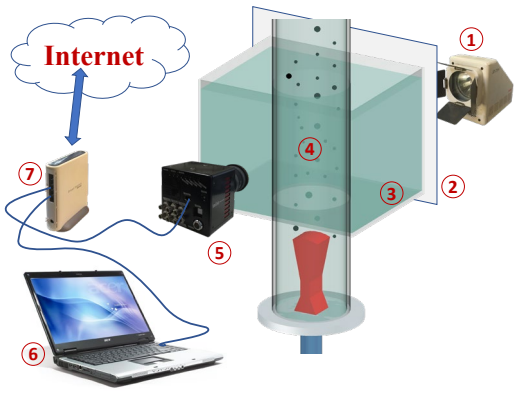


Fig. 2. Schematic diagram of the HSC measurement.
1– Light source; 2– Diffusion sheet; 3– Water box;
4– Scrubbing pool; 5– High-speed camera;
6– Computer; 7 – Router.

TABLE I. Setting Parameters of the High-Speed Camera.

Parameter:	Value:
Recording Rate	1000 fps
Shutter Speed	1/4000 sec.
Resolution	1024 x 1024 pixels
Number of frames	5457

the acrylic (1.49) and water (1.33), the refraction effect can be considered as insignificant.

The setting parameter of the high-speed camera is given in Table 1. By the limitation of the camera memory, a total of 5457 continuous frames were recorded. With the recording rate of 1000 frames per second, it permitted the measurement time gain to be 5.457 seconds.

C. *Application of the IoT*

For distance conduction of the experiments, an IoT network has been established for monitoring and controlling the experiments based on the three layers platform. The block diagram of the network is shown in Fig. 3. The IoT things consist of three switchbots, two smart valve controllers and one solenoid controlling valve which can be controlled via wifi internet connection. The measurement system includes a high-speed camera (Photron - Fastcam Mini AX50) which is controlled by a PC through an ethernet connection. This Lab-PC was also used to display the other parameters such as pressure, temperature, etc., which were collected from the sensors by a data logger device (Graphtech GL-200A) via USB connection. There were three web-cameras used to observe the performance of the apparatus.

The user can control the Lab-PC through the remote desktop control software (Chrome Remote Desktop Extension), therefore, control all the systems. The measurement data was automatically backed up into cloud storage. Therefore, it is easy to access and analyze by users from anywhere via an internet connection.

III. EXPERIMENTS

A. *Distance Measurement Schedule*

At first, the standby systems were turned on using smart switches and smart electric plugs. The water was filled into the pool using a smart valve controller at the feed water pipeline. The water level in the pool was controlled using both feed and drain water valves. The air compressor was turned on by a switchbot which was installed on the power button. The

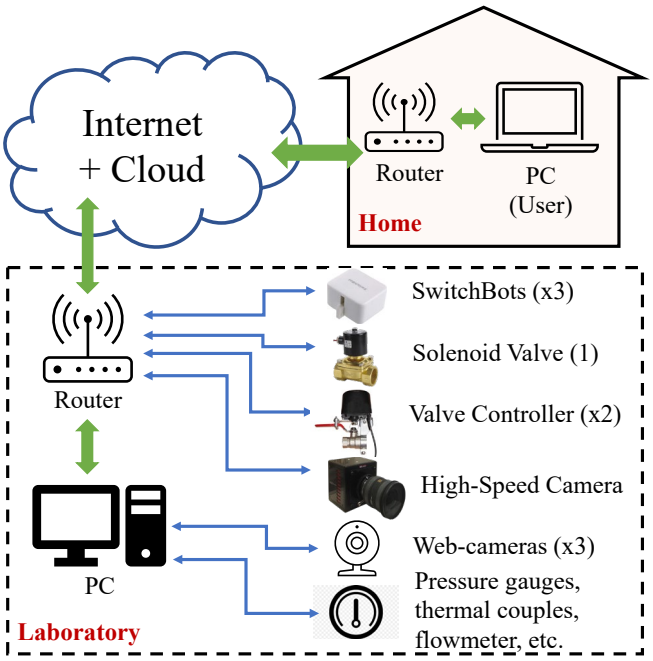


Fig. 3. Diagram of the monitoring and controlling system.

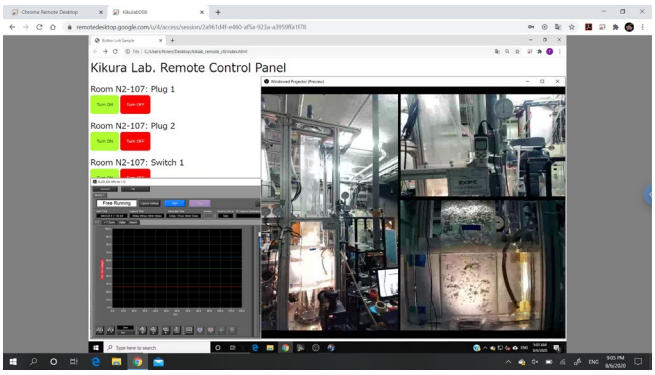


Fig. 4. Distance controlling of the apparatus.

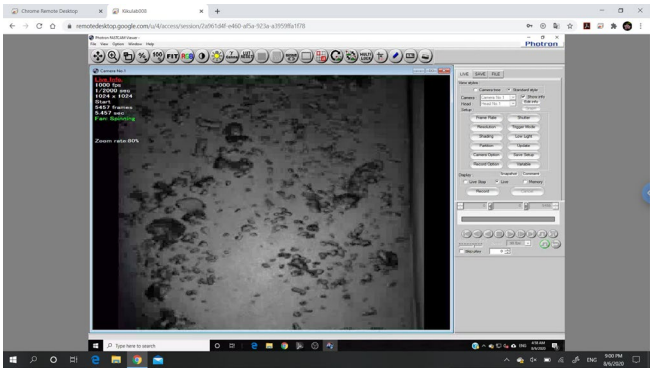


Fig. 5. Distance controlling of the HSC measurement.

air injected flow was controlled by the solenoid valve through PC's software. The high-speed camera, the backlight source, and the data logger device were also started using the switchbots.

Figures 4 and 5 show the interfaces of the apparatus and the HSC measurement controllers through the Chrome Remote Desktop Extension. The application is free and allows the user to control the lab PC from anywhere the internet is available.

The recorded images of the High-Speed Camera can be loaded into the home PC directly by backing-up data using cloud storage (Google Drive). Figures 6 and 7 show a raw image which visualized by the HSC for the air injected flow rate of 11.8 and 35.3 L/min, respectively. The flow pattern in the scrubbing pool of the wet-type filter was visualized at the height of 20 cm above the top of the Venturi Scrubber Nozzle (50 cm from the bottom of the Scrubbing Pool).

B. *Discussion*

The visualization of the two-phase flow of the FCVS is one of the most important measurements in the experimental research to improve the efficiency of an FCVS. However, the impact of the Covid-19 pandemic, the time for conduct experiments, has been limited. Therefore, the telemetric performance of experiments becomes important.

Although almost devices using in the experiment had no option for remote control. However, by adding several cheap, simple things, and free remote control software, the experiment can be remotely conducted.

Actually, the distance controlling of the system was a little difficult more than on-site controlling. It is because that, the

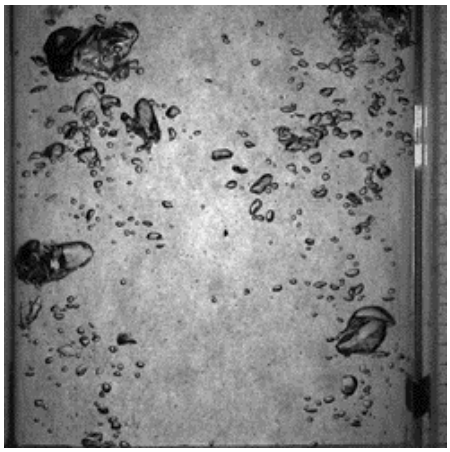


Fig. 6. Two-phase flow pattern in the pool. (Air inject flowrate of 11.8 L/min).

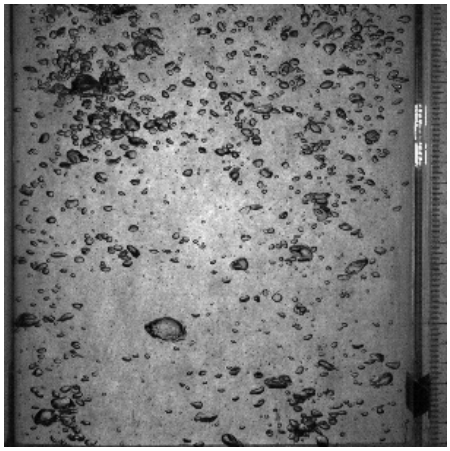


Fig. 7. Two-phase flow pattern of the pool. (Air inject flowrate of 35.3 L/min).

experience of adjustment of devices is different between on-site and off-site controlling. Sometime, the sensation of the operator is important. In addition, it takes the time to send the measurement data to the cloud, therefore, the time for each remote experiment is usually longer than on-site experiment. However, the experiment was conducted smoothly with the high-speed internet.

This experiment is an important part of our research project on the development of FCVS for the Nuclear Power Plant. The work is still in progress. The measurement is also the primary test for our collaborative research activities in the near future to construct the global telemetry research system, and for the projects on the decommissioning of the Fukushima Daiichi Nuclear Power Plant that the on-site measurements may not be conducted due to the high radiation.

IV. Conclusions

The IoT has been preliminarily applied to conduct a telemetric experiment. Such kind of tests should be more and more performed and developed. It is valuable in our future projects for the development of the remote laboratory and carrying out the telemetric studies, as well as increase the cooperation with other laboratories and researchers.

References

[1] R. O. Schlueter and R. P. Schmitz, “Filtered Vented Containments, Nuclear Engineering and Design,” vol. 120, pp. 93 - 103, 1990.

[2] T. Narabayashi, G. Chiba, A. Ishii, Y. Watanabe, N. Kitahara, K. Hirai, T. Masuda, Y. Nakasaka, N. Sato, K. Endo and T. Kobayashi, “Development of High Efficiency Multi-Nuclide Aerosol Filters for Radiation Protection during a Process of Cutting Core Debris for Fukushima Daiichi NPP,” Proc. 27th International Conference on Nuclear Engineering, Ibaraki, Japan, 2019.

[3] D. Jacquemain, “OECD/NEA/CSNI Status Report on Filtered Containment Venting,” OECD – Nuclear Energy Agency, Nuclear Safety NEA/CSNI/R7, 2014.

[4] M. Casini, D. Prattichizzo, and A. Vicino, “The automatic control telelab,” IEEE Control Syst. Mag., vol. 24, no. 3, pp. 36–44, Jun 2004.

[5] M. Leisenberg and M. Stepponat, “Internet of Things Remote Labs: Experiences with Data Analysis Experiments for Students Education,” 2019 IEEE Global Engineering Education Conference (EDUCON), Dubai, United Arab Emirates, pp. 22-27, 2019.

[6] T. S. El-Hasan, "Internet of Thing (IoT) Based Remote Labs in Engineering," 6th International Conference on Control, Decision and Information Technologies (CoDIT), Paris, France, 2019, pp. 976-982, 2019.

[7] M. V. Ramya, K. Purushothama and R. Prakash, “Design and Implementation of IoT Based Remote Laboratory for Sensor Experiments,” Int. J. Interactive Mobile Technologies, vol. 14, no. 9, pp. 227-238, 2020.

[8] M. Poongothai, A. L. Karupaiya and R. Priyadharshini, “Implementation of IoT based Smart Laboratory,” Int. J. Computer Applications, vol. 182, pp. 31-34, 2018.

Design of a low-cost NFC Door Lock for a Smart Home System

José Pacheco
Universidad Autónoma Metropolitana (UAM)
Dept. Metmáticas Aplicadas y Sistemas
Cuajimalpa, Ciudad de México
jpacheco.poo@gmail.com

Karen Miranda
Universidad Autónoma Metropolitana (UAM)
Dept. Sistemas de Información y Comunicaciones
Lerma, Estado de México
k.miranda@correo.ler.uam.mx

***Abstract*—The Internet of Things allows computers, people, and physical electronic devices to communicate with each other in all combinations, for example, people to people, people to computer, and computer to fridge by using a Internet connection. Hence, Smart home applications are very popular nowadays. The main goal of a smart home is to improve tasks such as home monitoring or security management. In particular, digital locks have evolved by using the Internet of Things technology. In this work, we present an NFC door lock implementation that is low-cost and user friendly based on Arduino boards as the microcontroller. The main advantage of our proposed prototype is that it can be easily installed and scalable on-demand due to the devices can work independently with each other.**

I. Introduction

Nowadays, wireless technologies' pervasive influence on our daily life is undeniable and, in this context, the Internet of Things (IoT) technology is in the spotlight. IoT allows our surrounding physical objects to be interconnected and exchange information via wireless networks such as Radio Frequency IDentification (RFID) and Wireless Sensor and Actuator Networks (WSAN) [1]. The main idea of IoT is that not only the objects interact with each other but also with the people in order to improve the quality of life [2].

Hence, Smart home technology is one of the main applications of IoT. The home, and in general building, automation aims to offer services like energy, security, healthiness, monitoring, accessibility, and communications management. Such management may be controlled from inside or outside the household; the latter by using an Internet connection and smart electronic devices, for example, to detect missing objects or avoid health risks.

In particular, security management is a wide study topic since humanity developed a need to keep privately their information and their belongings. In the latter case, by using mechanical devices known as locks [3]. For this reason, the locksmithing is in continuous growth and looking for creating more durable, reliable, and easy to manufacture locking mechanisms [4]. Naturally, locks evolved thanks to modern electronic devices and IoT technology as well. We can find a diversity of interfaces like PIN pad, touchpad, signature pad, or biometrics, used to authenticate the authorized persons [5]. Recently, ZigBee [6], Bluetooth [7], QR code [8], and RFID [9], [10] based proposals for door lock controllers have been presented as part of home automation systems.

978-1-7281-9615-2/20/$31.00

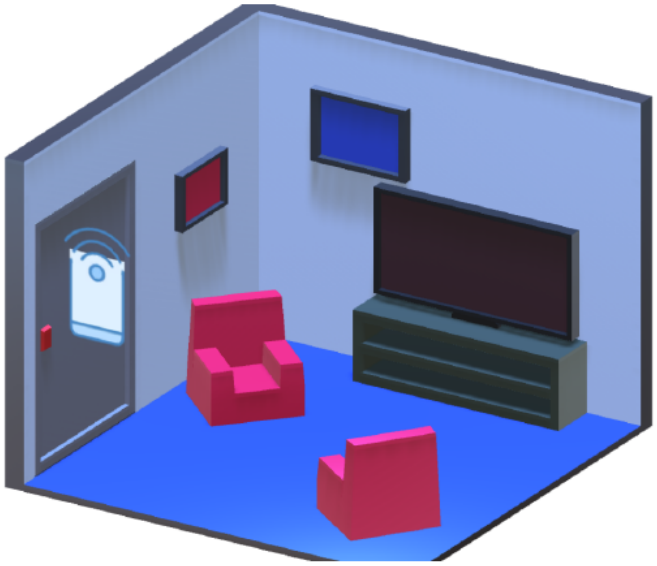


Fig. 1. Prototype diagram

In this paper, we present a low-cost door lock based on the Near-Field Communication (NFC) technology and the Arduino Mega board. Our prototype is easily installable and scalable in order to control access in a smart home as depicted in Figure 1. In this figure, we may appreciate a house that uses a digital lock at the main entrance. In our prototype, we use an NFC module as a reader and NFC card as a medium of authentication. Once the user is validated, the door opens automatically via the Arduino microcontroller.

The rest of this paper is structured as follows. In Section II, we provide a brief description of the proposed door lock hardware components and their specifications. In Section III, we describe the door lock overview and how the components and modules are integrated as well as the application interface. Thus, in Section IV we present the door lock implementation and some function examples. Finally, Section V concludes the paper.

II. Door Lock Components

A. Microcontroller

Arduino is a programmable electronic single-board microcontroller which is open-source in hardware and software. Currently, Arduino microcontrollers are wide use for prototyping and building digital devices due to is easily suitable for different applications.

Moreover, the Arduino board allows modules (called shields) to be integrated as interface hardware based on the Arduino's "easy to use" philosophy. Modules, for example, Bluetooth, Wi-Fi, cameras, microSD, GSM/GPRS, Joysticks, smoke detectors, LCD, or Ethernet may be integrated according to the problem.

Among the advantages of using an Arduino board as a microcontroller, we can find:

- Low cost. The Arduino UNO version can be purchased for less than $ 25 USD.
- Multi-platform. The Arduino IDE environment can be used on Windows, MacOS, and Linux.
- Size. Although the dimensions of the Arduino boards may vary, they are still very small components, approximately 5 x 6 x 1 cm.
- Shields. It is easy to integrate all types of sensors through Arduino's shields; Arduino MEGA has an analog (16) and a digital (54) sensors, for example.

Arduino boards are capable of reading sensors, buttons, and even Twitter messages and, accordingly, they generate an output, either by activating an engine, turning on a LED, or publishing content on the Internet. To perform these actions, the instructions are sent to the board's microcontroller by using the Arduino programming language and its Integrated Development Environment (IDE).

B. Reader and card

Near Field Communication (NFC) allows users to exchange information, by using NFC smart devices, with an intelligent object such as a tag, a reader, or an NFC-enabled mobile phone. Subsequently, the device decides what to do with the information received. An NFC reader can read from and write to an RFID tag, however, the NFC technology has more capabilities than the RFID one [11].

There are different operating modes for the NFC, such as reader/writer, peer-to-peer, and card emulation. Reader/writer mode allows the exchange of data between an NFC mobile and an NFC tag. Peer-to-peer mode allows the exchange of data between a pair of NFC mobiles. Finally, card emulation mode allows the user to interact with an NFC reader through their NFC mobile (such as a smart card). In particular, we are interested in the card emulation mode since one of the main application is access control.

NFC technology can be easily integrated into Arduino through a module. Although a mobile device is required to configure the NFC lock, it is possible to configure the NFC tags as cards for authentication. Since the NFC tags are inexpensive and they can be attached to almost any object thanks to the different NFC modes of operation mentioned above [12].

TABLE I
Arduino specifications

Microcontroller	ATmega2560
Operating Voltage	5V
Input Voltage (recommended)	7-12V
Input Voltage (limit)	6-20V
Digital I/O	Pins 54 (of which 15 provide PWM output)
Analog Input	Pins 16
DC Current per I/O	Pin 20 mA
DC Current for 3.3V	Pin 50 mA
Flash Memory	256 KB of which 8 KB used by bootloader
SRAM	8 KB
EEPROM	4 KB
Clock Speed	16 MHz
LED_BUILTIN	13
Length	101.52 mm
Width	53.3 mm
Weight	37 g

III. NFC door lock overview

In this work, we implement an NFC door lock that it should be low-cost, easy to install, and user friendly. In general, a typical door lock system consists of an automatic controller, an interface reader, and a back-end system [6], [9], [13]. Actually, the automatic controller is an actuator that acts depending on the input. Thus, our proposal presents those three main modules.

A. Arduino module

The Arduino project offers a broad range of products like starter kits that include tutorials for beginners such as Arduino UNO, Arduino 101, Arduino MICRO to mention a few. There are also products for projects that require better performance as Arduino MEGA or Arduino Zero.

In this work, we use Arduino Mega 2560 due to it is easily programmable and adaptable. Arduino Mega 2560 is a microcontroller based on the ATmega2560 with 54 digital input/output pins and 16 analog inputs, a USB port, and a power connector as specified in Table I.

B. NFC module

NFC is a short-range bidirectional wireless technology that uses a 13.56 MHz signal with a bandwidth of no more than 424 Kbps. The technology requires that two NFC devices be within a few centimeters of each other [12]. In particular, NFC Shield ITEAD PN532 is a module that may be integrated with Arduino and it is intended to read NFC devices and tags. The effective non-contact communication distance is three 3 cm as presented in Table II.

TABLE II
NFC SHIELD SPECIFICATIONS [14]

Size	40.5×43mm
IC	NXP PN532
Operating Voltage	3.3V
Power Supply Voltage	3.3 5.5V
Max Supply Current	150mA
Working Current(Standby Mode)	100mA
Working Current(Write Mode)	120mA
Working Current(Read Mode)	120mA
Indicator	PWR
Interface	SPI Interface, Std Raspberry Pi 20pins Interface

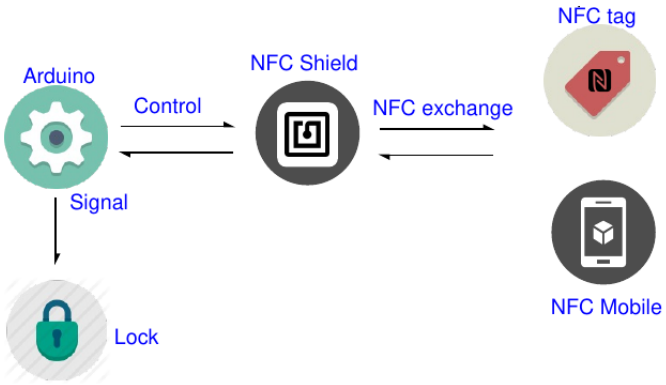


Fig. 2. Components interaction

C. Software module

We use Unity to develop an interface between the users and the NFC door lock. Unity is a cross-platform game engine. Despite that is a videogame oriented tool, it is used to create visualizations of virtual and augmented reality and also to create interactive applications in both two dimensions and three dimensions [15].

Unity 3D software uses programming routines to design, create, and develop interactive environments. Hence, we built an application to graphically feedback the user when he/she utilizes the NFC tag to access.

IV. NFC DOOR LOCK PROTOTYPE

We implement a prototype system for home automation based on a digital door lock with the NFC protocol. Figure 2 presents the system overview, where, we show the interaction of the components. On one hand, we have the Arduino board connected to the physical lock, the microcontroller will action the lock to unlock in case that is an authorized user. On the other hand, we have the NFC tag, inside a card. Finally, between both components, we have the NFC Shield, which reads the NFC tag and sends the gathered information to the Arduino board in order to verify the authentication of the user and act consequently.

Since the Arduino board could be provided with a memory medium, the database of authorized users can be stored on the board directly. Likewise, on the NFC tags, the unique identification number is stored. All the users are uniquely identified by a number, chosen randomly, sequentially, or other. Each time that a user wants to enter through the NCF lock should close the tag to the reader, then the number is processed at the microcontroller whether to allow access or not.

Regarding the power supply, Arduino Mega can be powered via a USB connection, AC/DC adapter, or battery. Thus, we assume that the power supply is not an issue for our implementation.

A. Modules Integration

In Figure 3, we may observe the prototype diagram. The figure presents the connection between the Arduino Mega

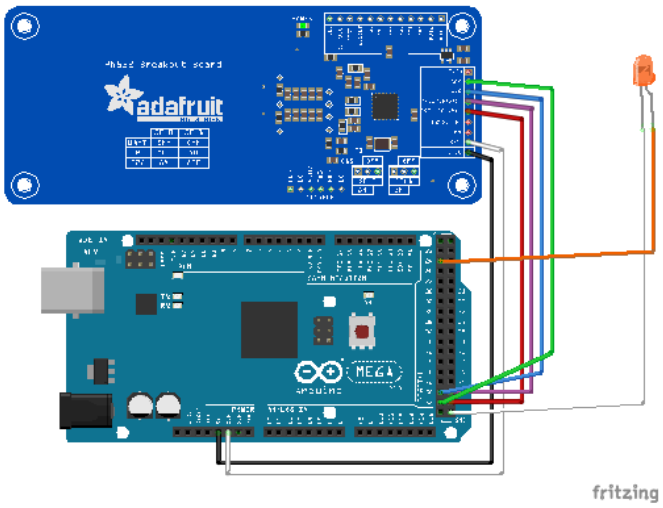


Fig. 3. Prototype diagram

module and the NFC Shield ITEAD PN532 module. We attach to digital pin 24 a LED light as a flag that the NFC module is connected correctly. We also connect both modules and the LED to the ground to prevent human contact with high voltages.

Thereafter, we connect the Clock signal from master to slave (SCK), Master Out Slave In (MOSI), Master In Slave Out (MISO), and Slave enabled signal, controlled by the master devices (NSS) communication protocols as part of the Serial Peripheral Interface (SPI) mode supported by Arduino's library for the NFC module.

Among the advantage of using SPI are, the protocol is simple as it uses a not complicated slave addressing, it is a fast protocol, compared to UART and I2C, data is transmitted continuously and bit by bit under the slave enable signal and shift pulse generated by the master device and it MISO and MOSI lines which means data can be transmitted and received at the same time.

Fig. 4. Homepage

Fig. 5. Example of an unauthorized user

B. Working example

In this section, we present the prototype implementation considering the scenario depicted in Figure 1, where the door lock depends on our prototype. Thereupon, Figure 4 shows the basic application homepage. In our proof of concept, we use two NFC cards. One with an authorized identification and the other with an unauthorized one. Moreover, our proposed module is connected to a screen that shows the interface created with Unity 3D for simulating the lock's panel.

Figure 5 depicts the moment when an unauthorized user attempts to unlock the door. In this example, we use an unauthorized NFC card. However, it is evident that if the user has not a card is not possible to unlock the door. As we may see at the screen, the lock is close, which means that the user is not allowed to enter.

Similarly, Figure 6 depicts the moment when an authorized user attempts to unlock the door. Contrary to the above example, we use an authorized NFC card and, therefore, the lock is open.

Fig. 6. Example of an authorized user

Finally, we show below an example of the Arduino code for getting information from the NFC module and the tag reading. Here, the microcontroller establishes the NFC module characteristics.

```
//Get information from device
Serial.print("Found chip PN5");
Serial.println((versiondata>>24)&0xFF,HEX);
Serial.print("Firmware ver.");
Serial.print((versiondata>>16)&0xFF, DEC);
Serial.print(' . ');
Serial.println((versiondata>>8&0xFF,DEC);
Serial.print("Supports ");
Serial.println(versiondata&0xFF, HEX);
//NFC tag reading
nfc.SAMConfig( );
```

V. Conclusion

In this paper, we presented a low-cost, easy installation, and scalable access control prototype implementation based on Arduino microcontrollers and Near-Field Communication technology. We have chosen the Arduino platform due to its flexibility to create hardware prototypes without worrying about the integration of complex electronic components as well as easily project tailored. Moreover, NFC technology is very promising and is gaining popularity. For example, some mobile phones use NFC technology to share power charges wirelessly, or they are used to pay as credit or debit cards. Hence, we believe that our prototype may function with the NFC mobile phone technology instead of tags.

Additionally, by using NFC technology, we consider that it is possible to improve lock security since it could be added to the native technology embedded in modern mobile phones such as facial or fingerprint recognition. So, we will study this as future work.

REFERENCES

[1] H. Chaouchi, "Chapter 1. Introduction to the Internet of Things," in *The Internet of Things: Connecting Objects to the Web*. Wiley and Sons, 2010, pp. 1–32.

[2] N. Mitton and D. Simplot-Ryl, "From the Internet of things to the Internet of the physical world," *Comptes Rendus Physique*, vol. 12, no. 7, pp. 669–674, 2011, Nanoscience and nanotechnologies: hopes and concerns.

[3] History of Keys, "History of Locks," accessed on July 2020. [Online]. Available: http://www.historyofkeys.com/locks-history/history-of-locks/

[4] The History of Keys, "History of Locksmithing," accessed on July 2020. [Online]. Available: http://www.historyofkeys.com/locks-history/history-of-locksmithing

[5] P. R. Nehete, J. P. Chaudhari, S. Pachpande, and K. P. Rane, "Literature Survey on Door Lock Security Systems," *International Journal of Computer Applications*, vol. 153, pp. 13–18, 2016.

[6] Y. T. Park, P. Sthapit, and J. Pyun, "Smart digital door lock for the home automation," in *TENCON 2009 - 2009 IEEE Region 10 Conference*, 2009, pp. 1–6.

[7] M. S. Hadis, E. Palantei, A. A. Ilham, and A. Hendra, "Design of smart lock system for doors with special features using bluetooth technology," in *Procedings of the International Conference on Information and Communications Technology (ICOIACT)*, Yogyakarta, Indonesia, Mar. 2018, pp. 396–400.

[8] A. F. M. Fauzi, N. N. Mohamed, H. Hashim, and M. A. Saleh, "Development of web-based smart security door using qr code system," in *2020 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*, 2020, pp. 13–17.

[9] G. K. Verma and P. Tripathi, "A Digital Security System with Door Lock System Using RFID Technology," *International Journal of Computer Applications*, vol. 5, pp. 6–8, 2010.

[10] K. Tshomo, K. Tshering, D. Gyeltshen, J. Yeshi, and K. Muramatsu, "Dual Door Lock System Using Radio-Frequency Identification and Fingerprint Recognition," in *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*, 2019, pp. 1–5.

[11] T. Igoe, D. Coleman, and B. Jepson, "Chapter 2. NFC and RFID," in *Beginning NFC*. O'Reilly Media, Inc., 2014, pp. 11–24.

[12] V. Coskun, K. Ok, and B. Ozdenizci, "Chapter 2. NFC and RFID," in *Professional NFC Application Development for Android*. Wrox, 2013, pp. 11–24.

[13] S. Nath, P. Banerjee, R. N. Biswas, S. K. Mitra, and M. K. Naskar, "Arduino based door unlocking system with real time control," in *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, 2016, pp. 358–362.

[14] Iteadstudio, "ITEAD PN532 NFC Module," accessed on July 2020. [Online]. Available: https://www.itead.cc/wiki/ITEAD_PN532_NFC_MODULE

[15] Unity, "Unity 3D," accessed on July 2020. [Online]. Available: http://unity3d.com

Light Spectra Optimization in Indoor Plant Growth for Internet of Things

Luzalen Marcos
Electrical, Computer and Biomedical Engineering
Ryerson University
Toronto, Canada
lmarcos@ryerson.ca

Kristiina Valter Mai
Electrical, Computer and Biomedical Engineering
Ryerson University
Toronto, Canada
kvmai@ryerson.ca

***Abstract*—This paper evaluates three plant grow light spectra suitable for indoor planting for Spinacia Oleracea (Spinach) and Lactuca Sativa (Lettuce). Indoor plant growing chambers and three commercial plant grow lights in the Red, Red-Blue and Yellow spectral range were used. Temperature, humidity and soil moisture were recorded. The spectral range of the grow light was found to affect the temperature, humidity and soil moisture, hence also affecting plant growth. Red light was found to be more suitable for spinach and red-blue light was found to be more suitable for lettuce. A system design allowing control of the spectral range by selection of the appropriate LED's based on the plant needs using the Internet of Things was developed. Future work will implement and test this feature together with other semi-automated plant growth systems such as irrigation and light intensity.**

Keywords—indoor plant growth, grow lights, light spectra, IoT

I. INTRODUCTION

Growing plants in general requires resources such as planting, maintaining soil temperature, watering the plants and others. Canadian farmers use different systems to ensure that plants are grown in the right temperature, moisture and pH of soil. During the whole life process of the plant, it is essential that plants are properly watered. The labours of the farmers are paid off after they harvest their crops. Although these responsibilities are the same with people who have plants indoors, there are additional factors that indoor plant growers need to consider such as light and how to power the system. However, people in Canada have a hectic lifestyle. With indoor planting gaining popularity in residential houses, it is essential to have a system that could assist in taking care of the plants.

This paper aims to provide an analysis of which type of light is better for Spinacia Oleracea, also known as Spinach, and Lactuca Sativa, also known as Lettuce. To achieve this, three commercial LED plant grow light were used to cultivate the plants placed in an indoor greenhouse. A solar powered system was designed and is being implemented to monitor soil moisture, temperature, humidity, light flux, and height detection of the plants. A review of related literature informed the design of the system.

This work was supported by the Natural Sciences and Engineering Research Council of Canada.

II. BACKGROUND

A. Plant growth requirements

Food production needs the right conditions, experience, and continued monitoring. Due to the Canadian climate, production is seasonal. To continuously produce crops throughout the year, greenhouse and indoor farming are used.

Advancement of technology has been applied in food production. For instance, [1] utilized internet of things (IoT) by implementing an automated farm irrigation system. Vij et al. [1] implemented a monitoring system to resolve irrigation problems of farmers. The system used a "wireless sensor network (WSN) to connect the components and machine learning system to determine irrigation based on crops and weather scenarios" [1]. Arduino Mega and Raspberry PI3 B+ were used as modules. The Arduino Mega has soil moisture sensor, PIR sensor, Temperature and humidity sensor, and MQ2 sensor for data collection and then, transmits data collected to the Raspberry PI. The Raspberry PI 3 B+ receives data via a WiFi Module, which is also used to transmit data into a cloud server. The data processing was done in the cloud server and then presented on a website. Multiple linear regression, Support Vector Regression (SVR) and Random Forest Regression were applied to predict soil moisture. For data presentation, a dynamic website was used to show real-time data to users. Their results showed that the Random Forest algorithm provides more accuracy than other algorithms. The authors suggested that proper installation of the setup should be done. Further, the system should be robust and dynamic.

Light emitting diodes (LED) are gaining popularity as a source of light for indoor plant growth to assist in photosynthesis. According to [2] lights with 400-700 nm help plants with photosynthesis. Thus, it is essential to determine which colour wavelengths and other factors to utilize in plant grow lights for indoor plant growth. There are different light colours (wavelengths) and plant purposes. One of the commonly used light colours is Red light. Red light, according to [2], contributes to the developing photosynthetic apparatus and shape of plants. Blue light aids in "stomatal opening, plant height, chlorophyll biosynthesis, seedling growth, root growth and during photosynthesis" [2]. The authors have also mentioned that blue light affects plant thickness. Lastly, Green light helped in plant growth and development. Moreover, different plants require different colour or wavelength. For instance, strawberries (*Fragaria ananassa*) require a

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

combination 660 nm red light and 450 nm blue LED lights [2]. In addition, distance of light source is important for indoor plant growth. If the light source is placed too close to the plants, it may burn the plants. On the contrary, if the light is placed farther from the plant, the plants may grow weak. To solve this, [2] used the inverse square law to determine the best distance between the light and plant. To elaborate, if the light source is 5 ft from the plants, then the plants will get 1/25 light from the source. Having some idea of which light to use for indoor plant growth can help ensure that the indoor plants will grow healthy.

Application of IoT and advanced technologies in indoor farming has been the focus of the research of Yang et al. [3]. Their proposed smart greenhouse system has three layers, namely, data sensing layer, cloud service layer, and user interaction layer. The data sensing layer collects data from different sensors. This includes an STM32-ARM processing chip, an ATK-ESP8266 WiFi module, a DHT11 for temperature and humidity sensing, an MH-410D for a CO2 infrared sensor, a ZE03-NO2 for an NO2 gas sensor, a ZE03-O3 for an O3 gas sensor, a ZE03-O2 for a O2 gas sensor, a BH170FVI for an ambient light intensity sensor and a video monitor of the plants. The data are then sent to the cloud service for data analysis and storage. After receiving data from the data sensing layer, the data are stored and organized using MySQL and Hadoop HDFS. Also, in the cloud data, joint analysis of indoor sensing data and plant growing data was done by using four algorithms, namely Support Vector Machines (SVM), Decision Tree Random Forests and Adaboost to predict growing status of plants. These algorithms were also compared by their accuracy, time and memory consumption. The results of the analysis can be viewed by the users in the user interaction layers via an Android App. Their result showed that indoor plant cultivation is made easier with IoT solutions. However, the authors suggested that analysis of big data is needed to utilize the information collected and the system should also consider user reference for the system. Yang et al. [3] also mentioned the need for collaboration between plant and technical experts to build a smart agricultural system.

Taking [1]–[3] as an inspiration, we present a system that could help assist indoor plant growers using IoT. Although our purpose is mainly focused on determining which light is more suitable for growing Spinach (Spinacia Oleracea) and Lettuce (Lactuca sativa}), our application of an automated greenhouse system using IoT is to determine the feasibility of having an intelligent monitoring system for indoor plant growth as mentioned as one of the aspects of [3]. Therefore, our system should determine the effects of different plant grow light spectra/colours in terms of temperature, humidity, light intensity, and soil moisture.

B. Plant Grow Lights

Plant grow lights are essential for indoor plant growth for they act as replacement of the Sun in assisting plant growth. Asiabanpour et al. [4] determined efficiency of optical fibre as indoor light for plant growth. Their study includes a simulation and actual testing of different sizes of optical fibres which were placed at different heights. For the simulation, AGI32, a light intensity design simulation software, was used to study light intensity. The study used one halogen lamp with 32,000 lux as light source. The light was passed through different optical fibres. Their simulation results showed that 20 mm fibre optical cables provide uniform light distribution and the empirical study showed that 20 mm fibre optic cable provided more light regardless of height. Although the experiment was focused on optimization of natural light, the actual effects of light on indoor plants was not examined.

There are different types of plant grow lights such as incandescent, fluorescent and Light Emitting Diode (LED). One of the early plant grow lights is the incandescent bulb, which has been traditionally used since it has a range of 600 to 800 nm spectrum. Incandescent light bulbs have been traditionally used for indoor planting. However, [5] found that incandescent lights for plant grow lights are not economical because they consume more energy than fluorescent plant grow lights and they have short service life. Therefore, fluorescent lights were used as artificial lighting for indoor plants. However, [5] mentioned that fluorescent light needs checking and replacement. Another type is Light Emitting Diode or LED. LED plant grow lights are known to have a longer life span [6]. Further, [6] mentioned that LEDs can produce specific wavelength, low thermal output and good light intensity.

These various light sources can be considered as a light source for indoor farming. For instance, Hurd [5] evaluated the effects of fluorescent and incandescent lamps on tomato plant growth in an indoor environment. Their result showed that far-red light is needed for tomato growth. Their results also showed that a warm-white fluorescent light bulb with an incandescent bulb assisted in tomato growth, but high blue light did not contribute to plant growth. Hurd [5] also mentioned that although incandescent bulbs contribute to a suitable wavelength to assist in plant growth, it is not economical to use incandescent bulbs for they require more energy.

Kohyama et al. [6] have suggested the use of Light Emitting Diodes (LED) as a replacement for incandescent bulbs without compensating the requirements for indoor plant growth. Their paper used two commercial LED plant grow lights and one incandescent bulb. One of the LED plant grow lights has a combination of Red and white diodes, the other LED has red, white and far-red diodes. They monitored the growth of different flowers: Ageratum (Ageratum houstonianum), Calibrachoa (Calibrachoa hybrida), Dianthus (Dianthus chinensis) and Petunia (Petunia hybrida) for "9-hour short day with or without 4-hour night interruption" [6]. Their results showed that Dianthus are shorter when they did not have enough light exposure. Additionally, they showed that different plants have different responses to different light colour.

Since LEDs are gaining potential as artificial lighting for indoor farming, we refer to [7] to determine the factors that were considered in designing and building a low-cost LED plant grow light. They have considered the type of chemical, temperature and wavelength of the plant grow device. In the end, the authors used a dual-emitting powder to fabricate a plant-grow light. The authors have suggested that a combination of red-blue light promotes plant growth both radially and vertically.

According to [8] and [9], different colours with different wavelength intervals have different effects on plant growth. For instance, red colour has a wavelength around 600 to 700 nm while blue has a wavelength between 400-500 nm. Even

though the combined lights have the same colour, their wavelength and effects on plants may vary as mentioned by [8], wherein they mentioned an experiment that compared the effects of different wavelengths of red colour on lettuce plants. The result showed that the biomass of lettuce was different under different wavelengths. Thus, it is essential to know which colour is beneficial for plant growth. In addition, different plants require different light intensity. Taking the result of [7], [10], [11] as an example, light intensity affects the growth and nutrients in plants. Their experiment used red-blue light to compare the photoperiod, which is the time period the plant receives light, and nutrients assimilated by Green and Red Lettuce.

Red light, according to [2], [7], [10] assists in "stem and radial growth, flowering and fruit production" while Blue light assists in vegetative growth. Combination of Red and Blue lights are commonly used for plant growth because they assist during photosynthesis [10]. With regards to photosynthesis, Green LED lights are also used. Yellow light is also used to replicate the Sun [12].

As mentioned before, different plants require different grow lights when considering indoor farming. Morales et al. [11] conducted research to determine which wavelength and photoperiod is best for Duckweed (Lemna gibba). Their LED grow light has a combination of red and blue diodes A digital lux meter and photosynthetic photon flux density was used to measure light intensity and wavelength, respectively. To measure temperature, the authors used a thermal imaging camera while calculation of digital lux was done using DIALux software, commonly used for lighting design. Their results showed growth and multiplication development under the LED plant grow light. The authors plan to use their designed LED plant grow light to optimize the "phytoremediation capability of Lemna gibba plants" [11].

Additionally, distance between the plant and the light is an important factor when doing indoor planting. Massa et al. [9] also used the Inverse Square Law to determine the appropriate height for the light. For example, if the light is 9 ft, then brightness will be 1/81.

For Green Leaf Lettuce (Lactuca sativa L. 'Lobjoits Green Cos') and Red leaf lettuce (Lactuca sativa 'Red Cos') plants, [10] determined the effects of light intensity or photoperiod on lettuce growth development. A modified LED lighting module that has blue, red and far red lights was used. For the first experiment, the effects of 100, 200, 300, 400 and 500 µmol m-2 s-1 photosynthetic photon flux density (PPFD) during an 18-hour photoperiod was determined while the effects of 12, 16, 18- and 24-hour photoperiod were measured. Leaf area, plant weight and plant heights were collected. Their results showed that 300-400 µmol m-2 s-1 PPFD and 16-18 hours photoperiod are needed for an effective nitrate assimilation on lettuce plants.

Sabzalian et al. [12] determined the effects of different LED lights in "Peppermint (Mentha Piperita), Spearmint (Mentha spicata) and Mentha Longfolia, Lentil (Lens culinaris), Basil (Ocimum basilicum) and four ornamentals" when placed in a field or greenhouse. A growth chamber was constructed and used for the experiment. For mint growth, the mint was placed in a chamber has a controlled temperature of 25 ± 2°C. The "green and potted flower" were grown in a greenhouse and a chamber. The results for the green and potted flower in different environments were compared. The results of [12] showed that LED lights are effective in plant growth and oil production from plants. The authors have also suggested that red-blue lights should be used to get the most benefits of LED lights.

From the history and review of LED plant grow lights of Yeh and Chung [13], they suggested that a combination of Green, Red and Blue lights are essential to have a full-spectrum of colour.

C. *Capacitive and Resistive Soil Moisture Sensors*

Adla et al. [14] compared four low-cost soil moisture sensors: two capacitive and two resistive soil moisture sensors, to determine which sensors can provide more precise and accurate results. Although the capacitive soil moisture sensors were not as cheap as the resistive soil moisture sensors, they outperformed the resistive soil moisture sensors in terms of accuracy and precision of data.

Aravind et al. [15] implemented a dual probe heat pulse (DPHP) that uses capacitive and resistive soil moisture sensors. Their research also did a comparison on both sensors based on accuracy, power and cost. By the end of their experiment, their results showed that in terms of cost, resistive sensors are cheaper. On the other hand, their designed DPHP sensor and capacitive soil moisture sensors provided better accuracy than the resistive soil moisture sensor. In terms of power, the power consumption of both resistive and capacitive sensors were considered negligible from their experiment while their DPHP required power.

III. Methodology

A. *Approach*

The system was designed to operate in a residential unit and be self sufficient between planting and harvest. Soil moisture, temperature and humidity data, irrigation measurement and plant images would be stored in a cloud-based server accessible by IoT monitoring so that the user

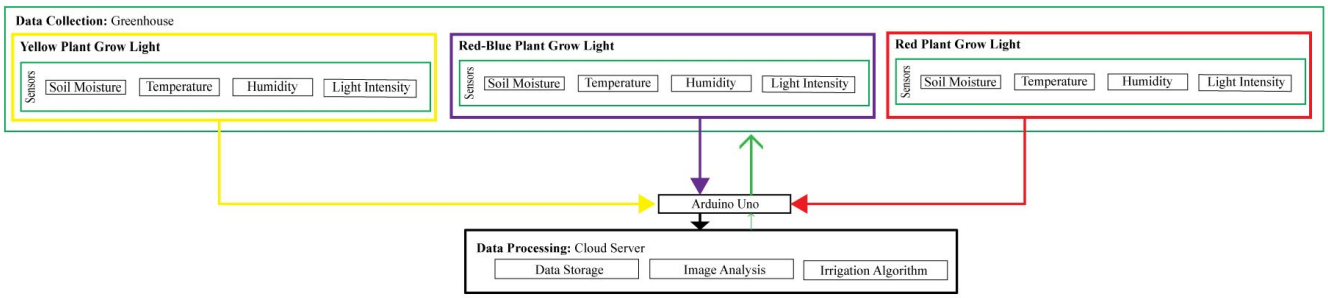


Figure 1: System Architecture

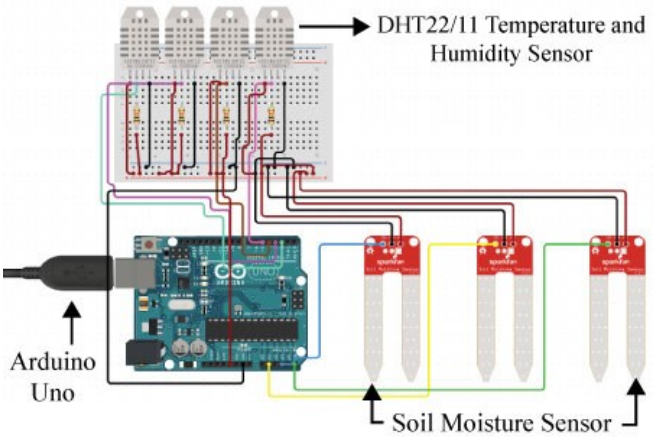


FIGURE 2: SCHEMATIC OF SENSORS AND MICROCONTROLLER

could check the status of the plants without having to be onsite. Furthermore, the system would be semi-autonomous, managing irrigation and light spectrum control automatically. The general system architecture is shown in Fig. 1.

B. System Architecture

As shown in Fig. 2, an Arduino Uno received data from the sensor. The specifications of the sensors are listed in Table 1 while the connection of the sensors to the microcontroller, using circuit.io, is shown in Fig. 2. From Fig. 2, there are four DHT11 sensors were used to monitor temperature and humidity of the three plant grow light chambers. The additional DHT11 sensor was placed outside the green house to monitor the temperature and humidity of the room. Further, from Table 1, two soil moisture sensors, namely, resistive and capacitive, were used to collect data to determine which soil moisture type will give more accurate data. Fig. 3 shows the algorithm that was used for data collection and plant irrigation.

TABLE 1: SENSOR SPECIFICATION

Sensor Type	VCC	Output Range
Resistive Soil Moisture	3.3-5V	0-550 Analog
Capacitive Soil Moisture	3.3-5V	0-850 Analog
Temperature Sensor	3.3-5V	0-50°C ± 2°C
Humidity Sensor	3.3-5V	20-80% ± 5%

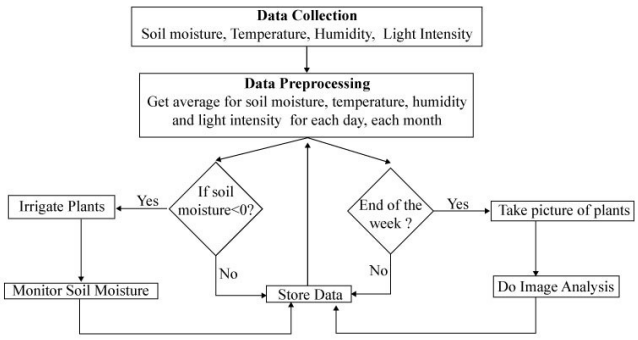


FIGURE 3: ALGORITHM OF THE MICROCONTROLLER

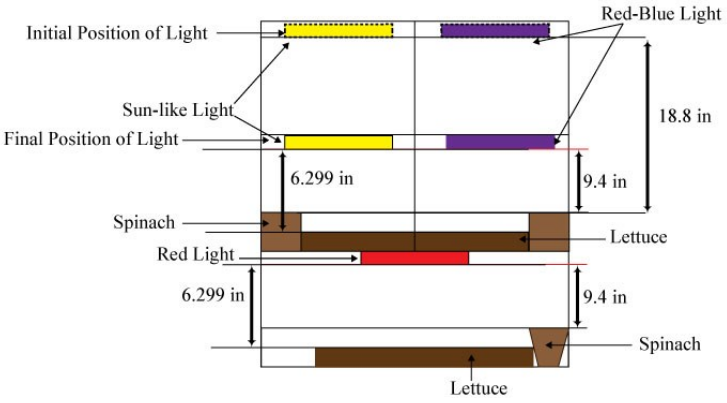


FIGURE 4: EXPERIMENTAL SETUP

C. Experimental Setup

A 27-in by 18-in by 63-in mini greenhouse and three commercial LED plant grow lights were used to cultivate the plants indoors. The specification of each plant grow light are shown in Table 2. The placement of the plants and plant grow lights were in divided chambers as shown in Fig. 4. As can be seen from Fig. 4, the distance between the spinach and plant grow light is 9.4-in while the lettuce has 6.2999-in distance from the plant grow light. Both the lettuce and spinach were grown from seeds. Images of the actual compartments in the greenhouse with the lights on are shown in Fig. 5.

TABLE 2: PLANT GROW LIGHT SPECIFICATION

Colour	Number of LEDs	Wattage	Wavelength (nm)
Yellow	169	75	380 – 800
Red-Blue	196	75	380 - 800
Red	80	40	Red: 660 Blue:460

D. Data Collection, Processing and Analysis.

Throughout the experiment, the plants were irrigated once a day and had the lights on for 12 hours per day. Soil moisture, temperature, humidity, and light intensity in flux of the plant chambers were collected and stored in a file for analysis and backup.

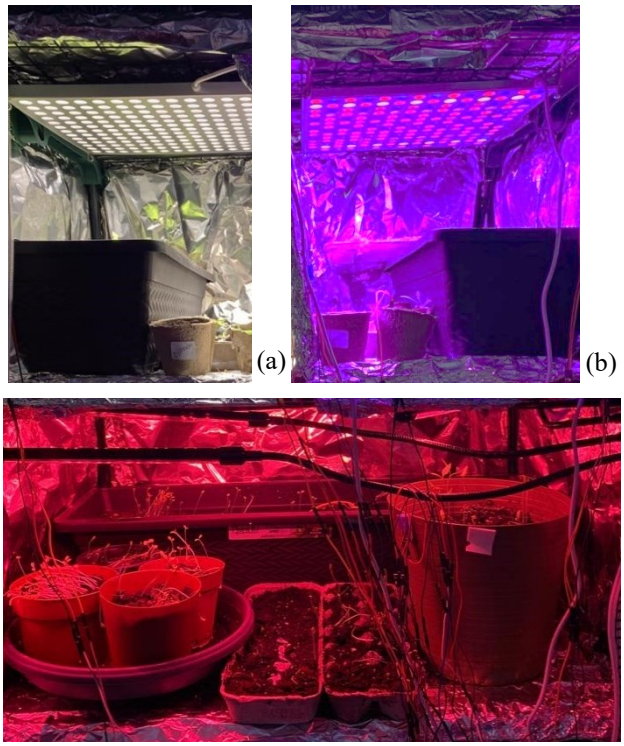


FIGURE 5: DIFFERENT COMPARTMENTS IN THE GREENHOUSE

IV. Results and Discussion

To determine which LED light colour is effective in spinach and lettuce growth, we used three different plant grow lights. Each plant grow light had its own compartment with temperature, humidity and soil moisture sensors. Both the spinach and lettuce plants grew in all the chambers with the system described in Section III.

The data obtained from the sensors show that different light colours have effects on temperature, humidity and soil moisture level. For instance, Fig. 6a shows that the red-blue plant grow light has the highest temperature compared to other plant grow lights on the first day, which may have been the result of the start-up characteristic. Further in Fig. 6a, b and c, on most days the yellow and red-blue light chambers have a higher temperature, humidity and soil moisture values than red plant grow light and room values. The resistive soil moisture sensor showed inconsistency in data collected. However, in Fig. 6c, the capacitive soil moisture sensor provided consistent data throughout the study.

With these results, the system design can be implemented with the findings of the suitable light spectra and intensities for growing plants such as spinach and lettuce in indoor growing conditions. The system will allow semi-autonomous operation, such that continuous human attention is not required, but the data is always accessible even when away from the residence and IoT allows for modification of system settings should a problem be found.

Plant growth is inherently variable as is to be expected with biological systems, and therefore periodic human intervention may be required. With the ideal of designing an energy efficient system, off-the-grid power can be added by harvesting excess light energy to power the sensors and data collection and transmission.

V. Conclusion

In conclusion, the following findings from this study will inform future IoT system design:

Plant grow light colour directly affects temperature and humidity of the compartment. We have observed interactions between different colour spectra and the other plant growing parameters.

Further, the results have given us an idea of how much data we need and how we can improve the system to obtain more accurate results. Therefore, these findings suggest the ability to operate a simple plant growth chamber with lighting spectra control.

VI. Future Work

During the time this paper is being written, an improved system is currently being implemented: ESP8266, a commonly used microcontroller for IoT applications, will replace the Arduino Uno microcontroller. Moreover, for the backend, Elasticsearch, Logstash and Kibana, also known as ELK stack, and Google Cloud will be used to provide better data analysis and visualization.

Future work will include the ability to modify the light spectrum by activating specific combinations of LEDs, controlling the irrigation based on sensor inputs and AI algorithms taking plant image inputs as controlling variables and user inputs to adjust the parameters. The system will be tested with various different plant species.

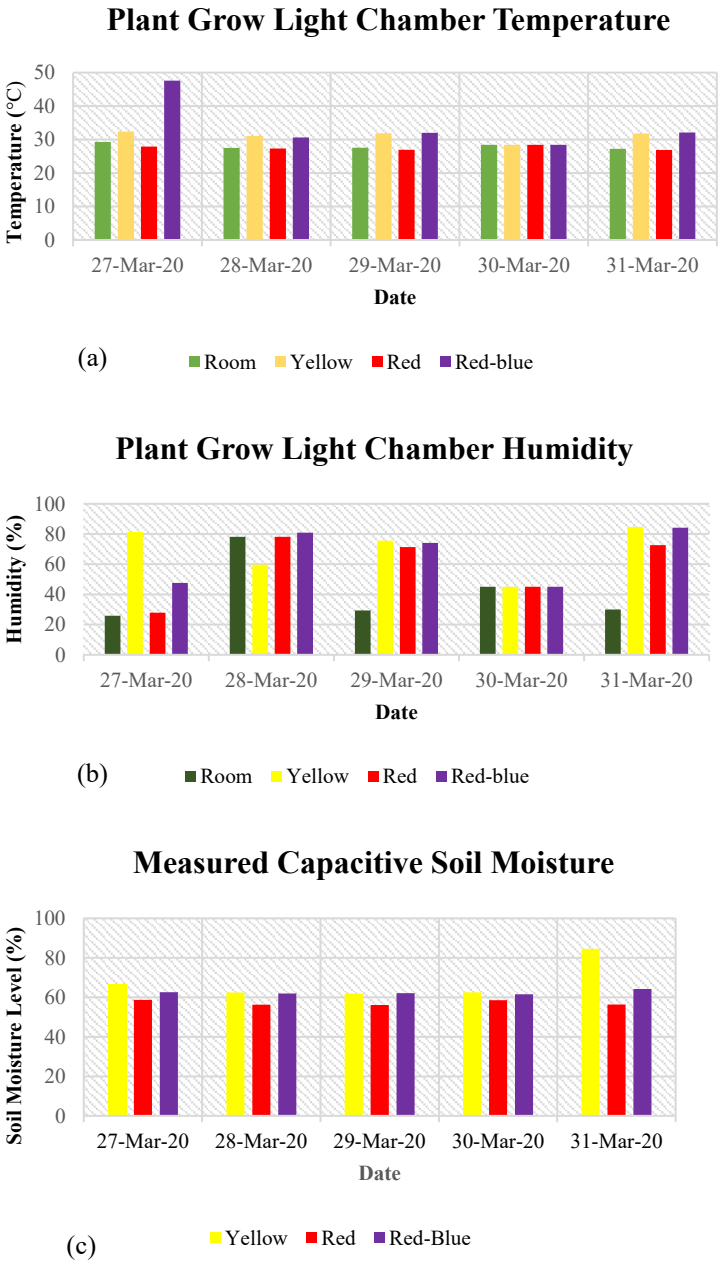


Figure 6: Measured Values

References

[1] A. Vij, S. Vijendra, A. Jain, S. Bajaj, A. Bassi, and A. Sharma, "IoT and Machine Learning Approaches for Automation of Farm Irrigation System," *Procedia Comput. Sci.*, vol. 167, pp. 1250–1257, 2020.

[2] M. Rehman, S. Ullah, Y. Bao, B. Wang, D. Peng, and L. Liu, "Light-emitting diodes: whether an efficient source of light for indoor plants?," *Environ. Sci. Pollut. Res.*, vol. 24, no. 32, pp. 24743–24752, 2017.

[3] J. Yang, M. Liu, J. Lu, Y. Miao, M. A. Hossain, and M. F. Alhamid, "Botanical Internet of Things: Toward Smart Indoor Farming by Connecting People, Plant, Data and Clouds," *Mob. Networks Appl.*, vol. 23, no. 2, pp. 188–202, 2018.

[4] B. Asiabanpour, A. Estrada, R. Ramirez, and M. S. Downey, "Optimizing Natural Light Distribution for Indoor Plant Growth Using PMMA Optical Fiber: Simulation and Empirical Study," *J. Renew. Energy*, pp. 1–10, 2018.

[5] R. G. Hurd, "The effect of an incandescent supplement on the growth of tomato plants in low light," *Annals of Botany*, vol. 38, no. 3. pp. 613–623, 1974.

[6] F. Kohyama, C. Wihitman, and E. S. Runkle, "Comparing flowering responses of long-day plants under incandescent and two commericial light-emitting diode lamps," *Prod. Mark. Reports*, vol. 24, pp. 1–24, August, 2011.

[7] L. Wang, H. Zhang, X. Zhou, Y. Liu, and B. Lei, "A dual-emitting core-shell carbon dot-silica-phosphor composite for LED plant grow light," *RSC Adv.*, vol. 7, no. 27, pp. 16662–16667, 2017.

[8] M. Olle and A. Viršile, "The effects of light-emitting diode lighting on greenhouse plant growth and quality," *Agric. Food Sci.*, vol. 22, no. 2, pp. 223–234, 2013.

[9] G. D. Massa, H. H. Kim, R. M. Wheeler, and C. A. Mitchell, "Plant productivity in response to LED lighting," *HortScience*, vol. 43, no. 7, pp. 1951–1956, 2008.

[10] A. Viršilė *et al.*, "Lighting intensity and photoperiod serves tailoring nitrate assimilation indices in red and green baby leaf lettuce," *J. Sci. Food Agric.*, vol. 99, no. 14, pp. 6608–6619, 2019.

[11] O. Morales, I. Gibran, A. B. Humberto, G. T. J. Fernando, P. T. A. Paola, and A. J. Oscar, "LED grow light for Lemna gibba plant," *2018 14th Int. Eng. Congr. CONIIN 2018*, pp. 1–4, 2018.

[12] M. R. Sabzalian *et al.*, "High performance of vegetables, flowers, and medicinal plants in a red-blue LED incubator for indoor plant production," *Agron. Sustain. Dev.*, vol. 34, no. 4, pp. 879–886, 2014.

[13] N. Yeh and J. P. Chung, "High-brightness LEDs-Energy efficient lighting sources and their potential in indoor plant cultivation," *Renew. Sustain. Energy Rev.*, vol. 13, no. 8, pp. 2175–2180, 2009.

[14] S. Adla, N. K. Rai, S. H. Karumanchi, S. Tripathi, M. Disse, and S. Pande, "Laboratory calibration and performance evaluation of low-cost capacitive and very low-cost resistive soil moisture sensors," *Sensors (Switzerland)*, vol. 20, no. 2, 2020.

[15] P. Aravind *et al.*, "A wireless multi-sensor system for soil moisture measurement," *2015 IEEE SENSORS - Proc.*, pp. 5–8, 2015.

Energy Scheduling and Computation Offloading for Building Operator using Parked Electric Vehicles

Shreya Bhardwaj, Shreyas Iyer, Tanvi Desai and Vamsi Krishna Tumuluru*
Dept. of Electronics and Communications Engineering
PES University, Bangalore
*Email: vamsikrishnatumuluru@pes.edu

***Abstract*—This paper proposes a new cooperative model between a building operator and the parked electric vehicles. The building operator owns few edge servers and must execute the delay sensitive computing requests generated in the building. The edge servers are powered by the grid and solar panels. Different from the traditional vehicular fog computing models where only computing services are provided by the electric vehicles, in the proposed model, the parked vehicle can also provide on-demand storage service or supply energy as a source. The building operator compensates the electric vehicles according to one of the services they offered during their parking. Compared to the price of electricity, the incentives paid to the electric vehicles for their service are assumed to be cheaper. The objective of the cooperative model is to minimize the power consumption costs of the edge servers operated by the building operator and the sum incentives paid to the parked electric vehicles for their service. The proposed cooperative model is formulated as a mixed integer program which can be solved by branch and bound solvers. Simulation results show the benefits of such cooperation to both the building operator and the parked electric vehicles under different scenarios.**

***Index Terms*—Edge servers, energy management, vehicles, computation offloading.**

I. Introduction

The idea of fog/edge computing is to provide low latency computing services to the clients within their proximity and reducing the dependence on remote datacenters [1]. Fog/edge computing deployments are becoming popular across a wide range of areas such as mobile communication, smart factory, office buildings, residential buildings, retail, vehicular networks, etc [2]–[4]. This growth can be attributed to the advent of 4G/5G networks and availability of small scale computing systems (4-10 servers) such as edge servers and micro datacenters [5]. The computing requests (e.g., delay sensitive tasks from mobile users, vehicles, sensors in buildings) can be offloaded to the edge servers due to their proximity. In case the edge servers are unable to process the requests (e.g., due to energy shortage or congestion due to workloads), then the workloads can be offloaded to the next set of fog servers in the hierarchy or to the cloud. Several existing works on fog/edge computing explored problems such as edge server placement, task offloading, energy management and radio resource allocation [2], [6]–[10]. In few of these works, the issue of energy management of edge servers was considered using strategies such as energy harvesting [2], [9] or switching the servers ON/OFF depending on the workloads [10].

978-1-7281-9615-2/20/$31.00 2020 IEEE

Recently, *vehicular fog computing* (VFC) has also emerged as an important approach for reducing the computation burden and energy consumption of the edge servers which may be owned by a mobile network operator or building operator [11]–[16]. In a VFC system, connected vehicles are capable of computing delay sensitive tasks using their limited on board computing units. Hence, some of the workloads can be offloaded to the connected vehicles thereby alleviating the energy requirements of the edge servers. The computing requests were generated by mobile users or sensors in buildings or other vehicles. The vehicles offering computing service could be either mobile or parked when the tasks are offloaded. The vehicles were compensated for their computing service either in the terms of monetary payments or through parking lots. In [11], workloads were generated by vehicles and offloaded by road side units (i.e., base station in vehicular networks) to other moving vehicles which could perform computing. The task assignment problem was formulated as an integer program and solved by particle swarm optimization. In [12], fixed edge servers (referred as fog nodes) attracted the vehicles with computing capabilities to take certain paths using pricing. The prices were based on the amount of requests in that region. The path selection and pricing problem was formulated as an integer program. In [13], a game-theoretic model was presented in which the mobile users could be served by either the edge servers of the base station or the vehicles. The amount of resources offered by the vehicles depended on the incentive paid by the base station. In [14], the mobile hot spots used parking lots to attract vehicles capable performing computing tasks. The parking lot reservation for the vehicles was decided through auctions depending on the computing resources and parking duration of the contending vehicles and computing requests. Vehicle to vehicle communication was used to allocate computing requests to moving vehicles in [15] and [16]. Reliability and task completion were the main concerns in these works as the vehicles were moving.

This paper proposes a cooperative model which enables edge servers (e.g., in a building) to execute the computation requests with the assistance of the parked electric vehicles. The power requirement of the edge servers is partly satisfied by the power bought from the grid and from pre-installed solar panels. Unlike existing works which only focused on the

on-board computing capabilities of the vehicles, the parked electric vehicles in the proposed work are assumed to offer one of the three types of services to the edge servers: a) they can act as on-demand energy storage devices (i.e., charge and discharge), b) they can act as energy providers (i.e., discharge power) or c) they can execute some of the requests offloaded by the edge servers. The building operator who manages the edge servers incentivizes the parked electric vehicles according to their service during their parking times. It should be noted that the proposed cooperative model allows greater flexibility to the building operator. The proposed cooperative model is formulated as an integer programming problem which is solved using branch and bound solver. The objective of the proposed problem is to minimize the total power cost incurred by the edge servers in the building and the sum incentives paid to the parked electric vehicles for their service.

The reminder of the paper is organized as follows. Section II presents the cooperative model for the parked electric vehicle and the building operator who runs the edge servers. The basis on which the incentives are paid to different types of service provided by the parked electric vehicles will be described. Section III provides the mathematical formulation which optimizes the power usage by the edge servers and the incentives for the parked electric vehicles. Section IV gives the results and the related discussion. Section V provides the concluding remarks.

II. System model

Consider a building consisting of a set of edge servers (denoted by $\mathcal{J}$) available for serving the delay-sensitive computing requests generated in the building (e.g., sensors, mobile users, surveillance cameras, etc.). The maximum number of requests that can be executed at edge server $j \in \mathcal{J}$ is denoted by M_j. Let A_t denote the number of requests generated in time slot $t \in \{1, \ldots, T\}$. The duration of each time slot is assumed to be 5 minutes. As shown in Fig. 1, the building is equipped with solar panels whose net output W_t is assumed to be known perfectly. The net output W_t represents the excess solar power available after deducting the fixed load of the building from the solar power output. The building operator draws power from the grid (denoted by decision variable p_t^g) at a price P_t and also relies on the solar power output to drive the edge servers. The power consumption of an edge server $j \in \mathcal{J}$ in time slot t, denoted by decision variable $p_{j,t}^e$, is given by

$$p_{j,t}^e = P_j^{idle} \cdot x_{j,t} + (P_j^{peak} - P_j^{idle}) \frac{m_{j,t}}{M_j} \tag{1}$$

In the above equation, P_j^{idle} and P_j^{peak} represent the idle power and peak power of server j respectively. Here, $x_{j,t}$ is a binary decision variable to represent the ON/OFF status of the edge server. Further, $m_{j,t}$ is a continuous decision variable to represent the number of requests being executed in time slot t by edge server $j \in \mathcal{J}$. Similar power consumption model for edge servers were also used in [6][1].

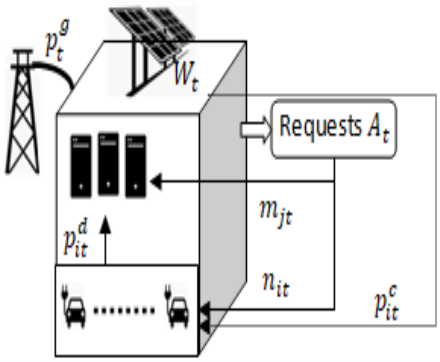


Fig. 1. System model

Additionally, the building operator has the option of using the services of the parked electric vehicles in the building to reduce the power consumption of its edge servers. The different types of services offered by the electric vehicles during their parking interval are: a) they can act as on-demand energy storage, b) they can act as energy sources or c) they can offer computing service. The building operator compensates the electric vehicles for their service using incentives denoted by αP_t (expressed in $\$/kWh$ where $\alpha \in [0, 1]$) and βP_t (expressed in $\$/kWh$ where $\beta \in [0, 1]$) for discharging and computing service respectively.

Let $\mathcal{I}$ denote the set of parking lots in the building. The set of arrival times and the set of departure times corresponding to each parking lot $i \in \mathcal{I}$ are denoted by $\mathcal{S}_i$ and $\mathcal{D}_i$ respectively. The sets $\mathcal{S}_i$ and $\mathcal{D}_i$ are assumed to be known from historical data. Based on $\mathcal{S}_i$ and $\mathcal{D}_i$, the set of parking times over a day, denoted by $\mathcal{H}_i$, is determined. Let $E_{i,t}^{in}$ denotes the energy available in parking lot i in slot $t \in \mathcal{S}_i$ due to an electric vehicle arrival. Each parking lot provides charging and discharging with an efficiency of η^c and η^d respectively. The maximum charging and discharging limits available in parking lot i are denoted by $\overline{C}_i$ and $\underline{C}_i$ respectively.

Let $\mathcal{R}_t^1$, $\mathcal{R}_t^2$ and $\mathcal{R}_t^3$ denote the set of parking locations in time slot t where parked vehicles offer on-demand storage, energy source and computing service respectively. Let $N_{i,t}$ denote the maximum number of requests that can be computed in parking lot $i \in \mathcal{R}_t^3$ in time slot $t \in \mathcal{H}_i$. The power spent by an electric vehicle due to computation (denoted by $p_{i,t}^v$ where $i \in \mathcal{R}_t^3$) is given by

$$p_{i,t}^v = P_i^{idle} \cdot y_{i,t} + (P_i^{peak} - P_i^{idle}) \frac{n_{i,t}}{N_{i,t}} \tag{2}$$

In the above equation, P_i^{idle} and P_i^{peak} represent the idle power and peak power corresponding to an electric vehicle in parking lot $i \in \mathcal{R}_t^3$ respectively. Here, $y_{i,t}$ is a binary decision variable to represent whether the building operator activated

[1]Other non-linear power consumption models can also be used. In such a case, they will be approximated using piecewise linear segments.

the computing service in parking lot $i \in \mathcal{R}_t^3$. Further, $n_{i,t}$ is a continuous decision variable to represent the number of requests being executed in time slot t at parking lot $i \in \mathcal{R}_t^3$. It should be noted that $p_{i,t}^v = 0$ for $i \in \mathcal{R}_t^1 \cup \mathcal{R}_t^2$.

Suppose an electric vehicle wants to act as an on-demand storage (i.e., $i \in \mathcal{R}_t^1$), then the building operator can charge and discharge the battery of that vehicle using the solar power available during its parking interval. The amount of energy stored in such vehicles at the time of their departure (denoted by $E_{i,t}^{out}$ where $i \in \mathcal{R}_t^1$ and $t \in \mathcal{D}_i$) must be same as the time of its arrival. Suppose an electric vehicle wants to act as a energy source (i.e., $i \in \mathcal{R}_t^2$), then the building operator is allowed to draw a maximum of $E_{i,t}^{dis}$ (expressed in kWh) from that vehicle during its parking interval. The value $E_{i,t}^{dis}$, where $i \in \mathcal{R}_t^2$, is specified by the vehicle owner. It is assumed that such electric vehicles were charged elsewhere at a price cheaper than the incentive paid by the building operator. Note that $E_{i,t}^{dis} = \infty$ for parked electric vehicles acting as on-demand storage (i.e., $i \in \mathcal{R}_t^2$). Further, $E_{i,t}^{dis} = 0$ if parked electric vehicle offers computing service only ($i \in \mathcal{R}_t^3$). Let $p_{i,t}^c$ and $p_{i,t}^d$ denote the power charged and power discharged in parking lot i in time slot t. It should be noted that $p_{i,t}^c = 0$ for $i \in \mathcal{R}_t^2 \cup \mathcal{R}_t^3$ (i.e., no charging is allowed for vehicles which act as a power source or offer computing service). Also note that $p_{i,t}^d = 0$ when $i \in \mathcal{R}_t^3$ (i.e., no power is discharged to the building from a vehicle offering computing service).

The above cooperation between the building operator and the different types of electric vehicles is formulated as a mixed integer programming problem (explained in Section III). The objective of the proposed cooperative model is to minimize the cost of power consumption of the edge servers owned by the building operator and sum incentives paid to the electric vehicles for their service.

III. Proposed Formulation

This section presents the mathematical formulation of the cooperative model for the building operator and the parked electric vehicles. The proposed formulation denoted by $\mathbb{P}1$ is given below. Here, Γ represents the set all decision variables used in the formulation. Tables I and II provide the list of all decision variables and the data used in the formulation respectively.

$$\mathbb{P}1: \min_{\Gamma} \sum_t P_t p_t^g + \alpha \sum_t \sum_i P_t p_{i,t}^d \eta^d + \beta \sum_t \sum_i P_t p_{i,t}^v \tag{3}$$

subject to: (1), (2), (4)-(18)

The first summation term in the objective function (3) represents the cost incurred by building operator in purchasing power for the edge servers from the grid. The second summation term in (3) represents the total incentive paid by the building operator to electric vehicles which acted as on-demand storage or energy sources (i.e., $i \in \mathcal{R}_t^1 \cup \mathcal{R}_t^2$) for drawing $p_{i,t}^d \eta^d$ amount of power. The third summation term in in (3) represents the total incentive paid by the building operator to electric vehicles which offered computing service (i.e., $i \in \mathcal{R}_t^3$) based on the power they spent performing computing (i.e., $p_{i,t}^v$).

Next, the constraints (4)-(17) are explained as follows. The sum of requests allocated to the edge servers ($\sum_j m_{j,t}$) and the sum of the requests allocated to the parked electric vehicles offering computing service (i.e., $\sum_{i \in \mathcal{R}_t^1} n_{i,t}$) must be equal to the number requests generated in the slot t (i.e., A_t). This condition is ensured by (4)

$$\sum_j m_{j,t} + \sum_{i \in \mathcal{R}_t^1} n_{i,t} = A_t \quad \forall t \tag{4}$$

TABLE I
List of Data and Description

Notation	Description
i, j, t	Indexes of parking lot, edge server and time slots
$\mathcal{I}$, $\mathcal{J}$	Set of parking lots and set of edge servers
$\mathcal{S}_i$	Set of arrival times in parking lot i
$\mathcal{D}_i$	Set of departure times from parking lot i
$\mathcal{H}_i$	Set of time slots when parking lot i is occupied
$E_{i,t}^{in}$	Incoming energy due to vehicle arrival in parking lot i in slot t
$E_{i,t}^{out}$	Energy in vehicle leaving parking lot i in slot t
$B_{i,t}$	Battery capacity available in parking lot i in slot t
$\overline{C}_i$, $\underline{C}_i$	Maximum charging and discharging limits in parking lot i
$\mathcal{R}_t^1$	Set of parking lots in t where on-demand storage is offered
$\mathcal{R}_t^2$	Set of parking lots in t where energy source is available
$\mathcal{R}_t^3$	Set of parking lots in t where computing service is available
$N_{i,t}$	Maximum requests handled by vehicle $i \in \mathcal{R}_t^3$
A_t	Number of requests arriving in slot t
M_j	Maximum requests handled by edge server j
W_t	Predicted net solar power output in slot t
P_t	Price of grid power in slot t
η^c, η^d	Charging and discharging efficiencies

TABLE II
List of Decision Variables and Description

Notation	Description
p_t^g	Power bought by building operator from grid
$p_{i,t}^d$	Power discharged in parking lot i in slot t
$p_{i,t}^c$	Power used for charging in parking lot i in slot t
$p_{i,t}^v$	Power spent in parking lot i in slot t for computation
$p_{j,t}^e$	Power required by edge server j in slot t
$n_{i,t}$	Number of request assigned to parking lot i in slot t
$m_{j,t}$	Number of request assigned to edge server j in slot t
$q_{i,t}$	Energy stored in parking lot i in slot t

Constraints (5) and (6) ensure that the number of requests assigned to the parked vehicles which offer computing service (i.e., $n_{i,t}$) and the number of requests assigned to the edge servers do not exceed their corresponding maximum request processing capacities respectively.

$$0 \leq n_{i,t} \leq N_{i,t} x_{i,t} \quad \forall i \in \mathcal{R}_3, \forall t \tag{5}$$

$$0 \leq m_{j,t} \leq M_j y_{j,t} \quad \forall j \in \mathcal{J}, \forall t \tag{6}$$

The power requirement of the edge servers in time slot t is satisfied based on the following: power bought from grid (p_t^g), solar power net output (W_t), power drawn from parked vehicles acting as on-demand storage ($p_{i,t}^d$ where $i \in \mathcal{R}_t^1$) and

power received from parked vehicles acting as power sources ($p^d_{i,t}\eta^d$ where $i \in \mathcal{R}^2_t$). This power balance condition given in (7) also includes the charging of vehicles used as on-demand storage $p^c_{i,t}$ (where $i \in \mathcal{R}^1_t$). However, as per (8) the charging is carried out only when solar power net output is available. Here, $p^c_{i,t}/\eta^c$ is the power supplied to the parked vehicle $i \in \mathcal{R}^1_t$ to actually store $p^c_{i,t}$ in that vehicle because of energy conversion losses.

$$\sum_j p^e_{j,t} + \sum_{i\in\mathcal{R}^1_t} p^c_{i,t} = p^g_t + W_t + \sum_{i\in\mathcal{R}^1_t\cup\mathcal{R}^2_t} p^d_{i,t}\eta_d \quad \forall t \tag{7}$$

$$\sum_{i\in\mathcal{R}^1_t} p^c_{i,t}/\eta^c \le W_t \quad \forall t \tag{8}$$

The evolution of stored energy in the parked electric vehicles is given by (9). Here, energy stored at the end of time slot t (i.e., $q_{i,t}$) depends on the energy stored in the time slot $t-1$ (i.e., $q_{i,t-1}$), the amount of energy added by charging (if $i \in \mathcal{R}^1_t$) and the amount of energy discharged (if $i \in \mathcal{R}^1_t \cup \mathcal{R}^2_t$) or utilized (if $i \in \mathcal{R}^3_t$). Note that the actual power discharged by parked vehicle $i \in \mathcal{R}^1_t \cup \mathcal{R}^2_t$ is $p^d_{i,t}$, however, only $p^d_{i,t}\eta^d$ is received by the building operator due to energy conversion losses.

$$q_{i,t} = q_{i,t-1} + p^c_{i,t} - p^v_{i,t} - p^d_{i,t} \quad \forall i, \forall t \in \mathcal{H}_i \setminus \mathcal{S}_i \tag{9}$$

The initial and final values of the stored energy are given by (10) and (11) corresponding to vehicle arrival and departure events respectively.

$$q_{i,t} = E^{in}_{i,t} \quad \forall i, \forall t \in \mathcal{S}_i \tag{10}$$

$$q_{i,t} = E^{out}_{i,t} \quad \forall i \in \mathcal{R}^1_t, \forall t \in \mathcal{D}_i \tag{11}$$

The total power discharged from a parked electric vehicle in parking lot $i \in \mathcal{R}^2_t$ in time slot $t \in \mathcal{H}_i$ must be exceed the maximum discharge allowed by that vehicle owner (i.e., $E^{dis}_{i,t}$). This condition is expressed in (12).

$$\sum_{k=\mathcal{S}_i(l)}^{t} p^d_{i,k} \le E^{dis}_{i,t} \quad \mathcal{S}_i(l)+1 \le t \le \mathcal{D}_i(l), l \in \{0,\dots,|\mathcal{S}_i|\} \tag{12}$$

In the above equation, $|\mathcal{S}_i|$ represents the number of arrival events that occur in parking lot i, $\mathcal{S}_i(l)$ represents the $l-th$ arrival event in parking lot i and $\mathcal{D}_i(l)$ represents the corresponding $l-th$ departure event from parking lot i. The left side of (12) represents the cumulative power discharged by the parked vehicle from the time slot it arrived to the time slot t. The right side of (12) represents the maximum allowed discharge by a parked vehicle.

The limits of charging and discharging in time slot t is given in (13) and (14) respectively.

$$0 \le p^c_{i,t} \le \overline{C}_i \quad \forall i, \forall t \tag{13}$$

$$0 \le p^d_{i,t} \le \underline{C}_i \quad \forall i, \forall t \tag{14}$$

The amount of energy stored in a parked vehicle should not exceed the available battery capacity (denoted by $B_{i,t}$). This condition is given by

$$0 \le q_{i,t} \le B_{i,t} \quad \forall i, \forall t \tag{15}$$

The non-negativity conditions for the continuous decision variables are expressed below.

$$p^v_{i,t}, p^c_{i,t}, p^d_{i,t} \ge 0 \quad \forall i, \forall t \tag{16}$$

$$p^e_{j,t} \ge 0 \quad \forall j, \forall t \tag{17}$$

$$p^g_t \ge 0 \quad \forall t \tag{18}$$

The above formulation is a mixed integer programming problem which is solved using branch and bound solver (e.g., CPLEX).

IV. Results and Discussion

For the purpose of simulation, a building with 5 edge servers and 12 parking lots is considered. Each edge server can process up to $M_j = 6$ requests per slot. The mean rate of arrival of the computing requests is chosen in the range of $[10, 20]$ requests per slot. Here, each slot $t \in \{1,\dots,96\}$ is assumed to have a duration of 15 minutes. The solar capacity was set to 10 kW and the solar power output per slot were obtained based on the normalized data given in [17]. The arrival and departure times for the electric vehicles were generated using the normal distributions $\mathcal{N}(36,5)$ and $\mathcal{N}(68,5)$ respectively. The arriving vehicles were assigned parking lots randomly and the sets associated with each parking lot $i \in \{1,\dots,12\}$ were updated accordingly. The initial stored energy $E^{inc}_{i,t}$ was randomly generated from $[4, 10]$ kWh. The battery capacity of all vehicles was set to 16 kWh. The maximum charging limit $\overline{C}_{i,t}$ and discharging limit $\underline{C}_{i,t}$ were set to 1 kW. The charging and discharging efficiencies were set to 0.9. For parked vehicles offering computing service, the maximum requests that can be handled per slot ($N_{j,t}$ where $j \in \mathcal{R}^3_t$) was set to 2. The maximum allowed discharge for a parked vehicle offering power supply (i.e., $E^{dis}_{i,t}$ where $j \in \mathcal{R}^2_t$) was set to 2 kW. The retail price P_t was set to 5 rupees-per-kWh whereas the incentive factors α and β were set to 0.6.

Fig.2 shows the effect of the request rate on the total cost incurred by the building operator (i.e., cost of grid power to run the edge servers and total incentive to the parked vehicles). For this simulation, equal number of parked electric vehicles of each type were chosen. It can be seen that as arrival rate increases, the total cost increases. However, the cost incurred due to the cooperation with the parked vehicles is lower compared to the case without such cooperation. This is evident from Fig. 3 which shows that the grid power bought by the building operator is lesser in the case with cooperation than that without cooperation.

Next, the effect of the number of parked vehicles on the total cost incurred by the building operator will be examined. For these simulations, vehicles offering the same type of service are considered to be in the system. Here, the request rate λ was set to 20.

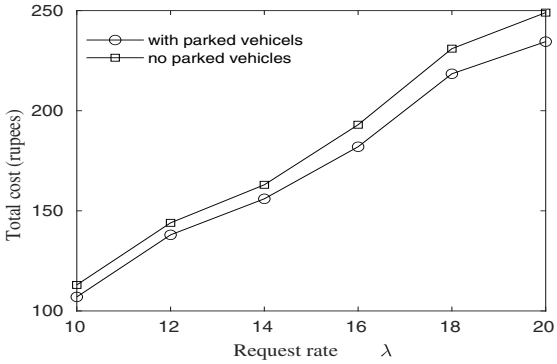


Fig. 2. Effect of request rate on total costs.

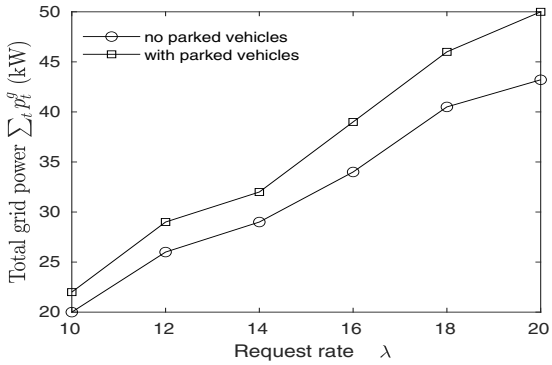


Fig. 3. Effect of request rate on total power bought from the grid.

Fig. 4 shows the effect on the total cost due to the number of parked vehicles which offer on-demand energy storage (i.e., $i \in \mathcal{R}_t^1$ where $i \in \mathcal{I}$). It can be observed that as the number of such parked vehicles increases, the total cost decreases. It should be noted that the parking intervals of these vehicles coincided with the peak hours of solar power. It can be noted that beyond 6 such parked vehicles, there was on improvement in the total cost. This is because the maximum possible power harvested with the assumed solar capacity was stored and retrieved by just 6 vehicles. Fig. 5 shows the amount of power stored and retrieved from the parked electric vehicles acting as on-demand energy storage units. As the number of such parked vehicles increased, the opportunities for reusing the harvested power increased thereby reducing the grid power purchased from retail.

Fig. 6 shows the effect on the total cost due to the number of parked vehicles which act as energy sources (i.e., $i \in \mathcal{R}_t^2$ where $i \in \mathcal{I}$). It can be observed that as the number of such parked vehicles increases, the total cost decreases. It can also be observed that such vehicles played a greater role especially when the solar power generation is not present. Fig. 7 shows as these parked electric vehicles increased, the building operator bought lesser grid power from the retail. Further, it can be seen that the building operator benefited as the parked vehicles agreed to allot more power for discharging (i.e., maximum allowed discharging over the parked interval E_i^{dis}).

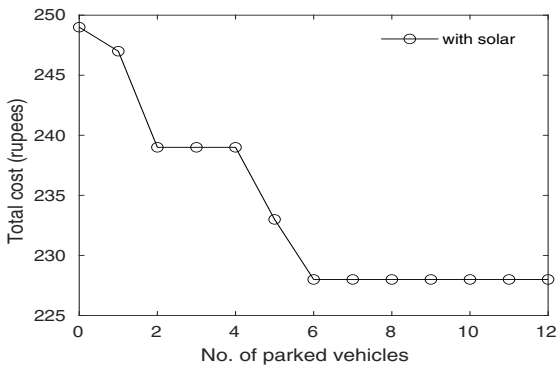


Fig. 4. Effect of number of energy storage vehicles on total cost of the building operator.

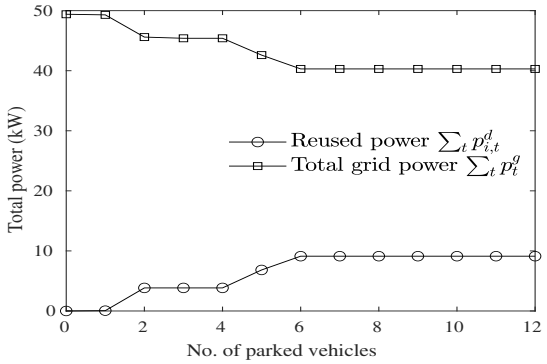


Fig. 5. Effect of number of energy storage vehicles on the power consumption behavior.

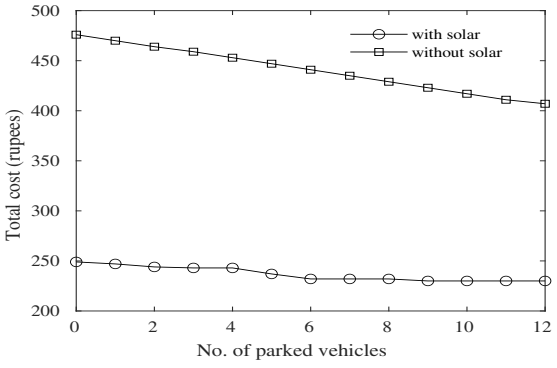


Fig. 6. Effect of number of power supplying vehicles on total cost of the building operator.

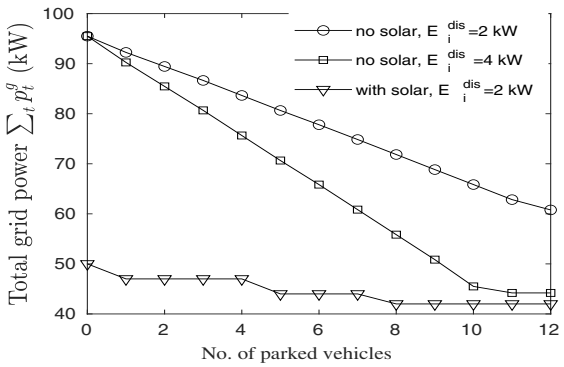


Fig. 7. Effect of number of power supplying vehicles on the power consumption behavior.

Fig. 8 shows the effect on the total cost due to the number of parked vehicles which offer computing service (i.e., $i \in \mathcal{R}_t^3$ where $i \in \mathcal{I}$). It can be seen that as the number of such vehicles increases, the total cost decreases. It is observed that such vehicles played a greater role when there was no solar capacity in the building. The total number of requests processed by the vehicles (i.e., $\sum_i \sum_t n_{i,t}$) increased with the number of vehicles. When the solar capacity was 10 kW, $\sum_i \sum_t n_{i,t}$ increased from 16 to 69 requests as number of vehicles increased from 1 to 12. When no solar capacity was considered, $\sum_i \sum_t n_{i,t}$ increased from 82 to 710 requests as number of vehicles increased from 2 to 12.

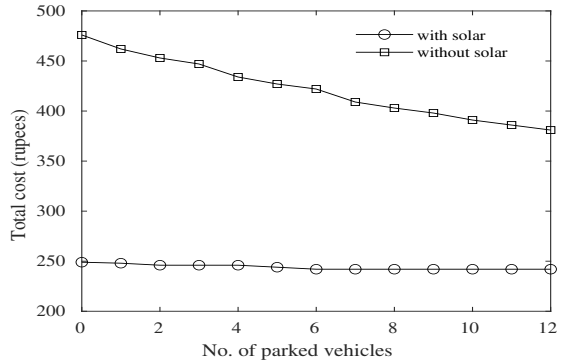


Fig. 8. Effect of number of computing vehicles on total cost of the building operator.

V. Conclusion

This paper proposed a new cooperative model for parked electric vehicles and a building operator who supports delay-tolerant computing requests using edge servers. The objective of the cooperative model was to minimize the power consumption cost of edge servers run by the building operator while duly compensating the parked electric vehicles according to their service. Unlike the traditional vehicular fog computing models which only considered computation offloading to electric vehicles, the proposed model also bought power from electric vehicles and used some vehicles as energy storage units. The results showed that the building operator power consumption cost associated with the edge servers reduced significantly when the electric vehicles of different types were involved. It was observed that the vehicles acting as on-demand storage units enhanced the performance when solar power capacity was pre-installed. Vehicles which offered computing service and which offered power supply were very useful under all circumstances (i.e., with or without pre-installed solar power capacity). Future work will focus on uncertainty in the data and optimal solar capacity estimation.

References

[1] M. Chiang, "Fog networking: An overview on research opportunities," *CoRR*, vol. abs/1601.00835, 2016. [Online]. Available: http://arxiv.org/abs/1601.00835

[2] F. Guo, L. Ma, H. Zhang, H. Ji, and X. Li, "Joint load management and resource allocation in the energy harvesting powered small cell networks with mobile edge computing," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2018, pp. 299–304.

[3] L. Chen, P. Zhou, L. Gao, and J. Xu, "Adaptive fog configuration for the industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 10, pp. 4656–4664, 2018.

[4] A. Kovalenko, R. F. Hussain, O. Semiari, and M. A. Salehi, "Robust resource allocation using edge computing for vehicle to infrastructure (v2i) networks," in *2019 IEEE 3rd International Conference on Fog and Edge Computing (ICFEC)*, 2019, pp. 1–6.

[5] V. Avelar, "Cost benefit analysis of edge micro data center deployments," Schneider Electric, May 2017, white Paper 223.

[6] Y. Li and S. Wang, "An energy-aware edge server placement algorithm in mobile edge computing," in *2018 IEEE International Conference on Edge Computing (EDGE)*, 2018, pp. 66–73.

[7] G. Manasvi, A. Chakraborty, and B. S. Manoj, "Social network aware dynamic edge server placement for next-generation cellular networks," in *2020 International Conference on COMmunication Systems NETworkS (COMSNETS)*, 2020, pp. 499–502.

[8] R. Bruschi, F. Davoli, C. Lombardo, and O. R. Sanchez, "Evaluating the impact of micro-data center (dc) placement in an urban environment," in *2018 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN)*, 2018, pp. 1–7.

[9] G. Zhang, Y. Chen, Z. Shen, and L. Wang, "Energy management for multi-user mobile-edge computing systems with energy harvesting devices and qos constraints," in *2018 27th International Conference on Computer Communication and Networks (ICCCN)*, 2018, pp. 1–6.

[10] M. Daraghmeh, I. Al Ridhawi, M. Aloqaily, Y. Jararweh, and A. Agarwal, "A power management approach to reduce energy consumption for edge computing servers," in *2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC)*, 2019, pp. 259–264.

[11] C. Zhu, J. Tao, G. Pastor, Y. Xiao, Y. Ji, Q. Zhou, Y. Li, and A. Yl-Jski, "Folo: Latency and quality optimized task allocation in vehicular fog computing," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4150–4161, 2019.

[12] S. Liao, J. Li, J. Wu, W. Yang, and Z. Guan, "Fog-enabled vehicle as a service for computing geographical migration in smart cities," *IEEE Access*, vol. 7, pp. 8726–8736, 2019.

[13] Z. Zhou, P. Liu, J. Feng, Y. Zhang, S. Mumtaz, and J. Rodriguez, "Computation resource allocation and task assignment optimization in vehicular fog computing: A contract-matching approach," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 3113–3125, 2019.

[14] Y. Zhang, C. Wang, and H. Wei, "Parking reservation auction for parked vehicle assistance in vehicular fog computing," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 3126–3139, 2019.

[15] S. S. Shah, M. Ali, A. W. Malik, M. A. Khan, and S. D. Ravana, "vfog: A vehicle-assisted computing framework for delay-sensitive applications in smart cities," *IEEE Access*, vol. 7, pp. 34 900–34 909, 2019.

[16] Q. Wu, H. Ge, H. Liu, Q. Fan, Z. Li, and Z. Wang, "A task offloading scheme in vehicular fog and cloud computing system," *IEEE Access*, vol. 8, pp. 1173–1184, 2020.

[17] "Ontario Independent Electricity System Operator: Power data," Jan 2019. [Online]. Available: http://www.ieso.ca/en/power-data?chart=supply

Ichor - an IoT, Cloud, and Mobile Technologies Based Noninvasive Glucose Monitoring System

Satvik Dasari
Westwood High School
Round Rock ISD
Austin, USA
srdasari2@gmail.com

***Abstract*—According to the Centers for Disease Control and Prevention, 34.2 million Americans have diabetes, which is 10.5% of the US population. Additionally, 88 million or 34.5% of the US adult population have prediabetes. People with diabetes are required to monitor their sugar levels periodically to avoid hypoglycemia (low blood sugar) or hyperglycemia (high blood sugar). The most common method to measure glucose levels is invasive and involves pricking fingers. Although there are noninvasive methods, they are very expensive. This paper proposes an innovative noninvasive system to measure glucose levels in blood using sensors, a microcontroller, Cloud, and mobile technology. Mechanisms built into this device send alerts to caregivers if glucose levels are unsafe. The glucose monitoring system is also integrated into an Internet of Things (IoT) analytics platform that shows continuous and historical data. The solution is low-cost, easy to use, and allows continuous monitoring. The prototype was validated using a dextrose solution and real-time readings. The results show a close relationship between actual glucose levels and the data collected from the noninvasive glucometer.**

Keywords—Glucose, Noninvasive, Diabetes, IoT, Cloud, Mobile App

I. INTRODUCTION

According to the World Health Organization (WHO), more than 422 million adults have diabetes globally [1]. As per the report published by the Center for Disease Control and Prevention (CDC), in the United States alone, there are 34.2 million diabetics, and 88 million people have prediabetes [2]. Worldwide, there were over 1.6 million deaths caused by diabetes in 2016. Higher than optimal glucose levels caused an additional 2.2 million deaths due to an increased risk of cardiovascular and other diseases in 2012 [1]. People with diabetes are required to monitor their sugar levels constantly to avoid hypoglycemia (low blood sugar) or hyperglycemia (high blood sugar) [3]. The most common solution to monitor glucose levels is an invasive glucometer. Current glucometers involve pricking a finger and dripping blood on to a test strip. They produce accurate results, but they are painful, risk infection, and are expensive. Furthermore, they are not suitable for continuous monitoring. Although there are some noninvasive glucometers on the market, they are very expensive. For example, the noninvasive glucometer from a certain company costs around $2100, and another company offers a 14 day sensor kit for $83.

The goal of Ichor is to design, build, and evaluate a noninvasive glucometer using sensors, IoT, Cloud, and mobile technologies. In addition, the system should be low-cost, scalable, accurate, reliable, and provide remote monitoring and maintenance capabilities.

The summary of contributions include:

- a noninvasive glucometer using sensors and an Internet of Things (IoT) device
- a firmware that processes the data from sensors and provides local alerts using Red-Yellow-Green LEDs
- an Android mobile app that shows glucose levels and color-coded alerts so that caregivers can monitor glucose levels remotely
- integration of the glucose monitor device with IoT analytics platforms such as ThingSpeak for storing and displaying historical data
- a mechanism to monitor and maintain glucometers remotely

The rest of the paper is organized as follows. Section II describes the system architecture of Ichor. Section III describes the mathematical model used by this solution. Section IV details the experiments and data collection. Section V provides an evaluation and analysis of Ichor. Section VI concludes the paper.

II. SYSTEM ARCHITECTURE

Ichor consists of five main components:

- the glucometer device
- the Cloud
- the glucometer console
- the alert mechanisms (Mobile App and local LEDs)
- the IoT analytics

Fig. 1 shows the overall architecture of Ichor. The Glucometer device consists of a near-infrared (NIR) light [11], a NIR photodetector, a noise filter, an amplifier, and a microcontroller. A LED1550E from Thorlabs is used as the NIR light, and it emits light with a spectral output centered at around 1550 nm [5]. The corresponding photodiode from Thorlabs (FGA10) was used because of its peak responsivity at 1550 nm [6] [8]. A noise filter was built using a resistor and a capacitor [10]. An LM358 and resistors were used to build an amplifier. A microcontroller from Particle reads the data from the sensor, processes it, and turns on corresponding LEDs (red, yellow, and green) to indicate glucose levels [7]. Additionally, the microcontroller also has a 3G module, which helps propagate the data to the Cloud. Also, if 3G is not required, the microcontroller comes in a version that can send data via Wi-Fi. Fig. 2 shows the detailed flow of the firmware programmed onto each one of the glucometer devices.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

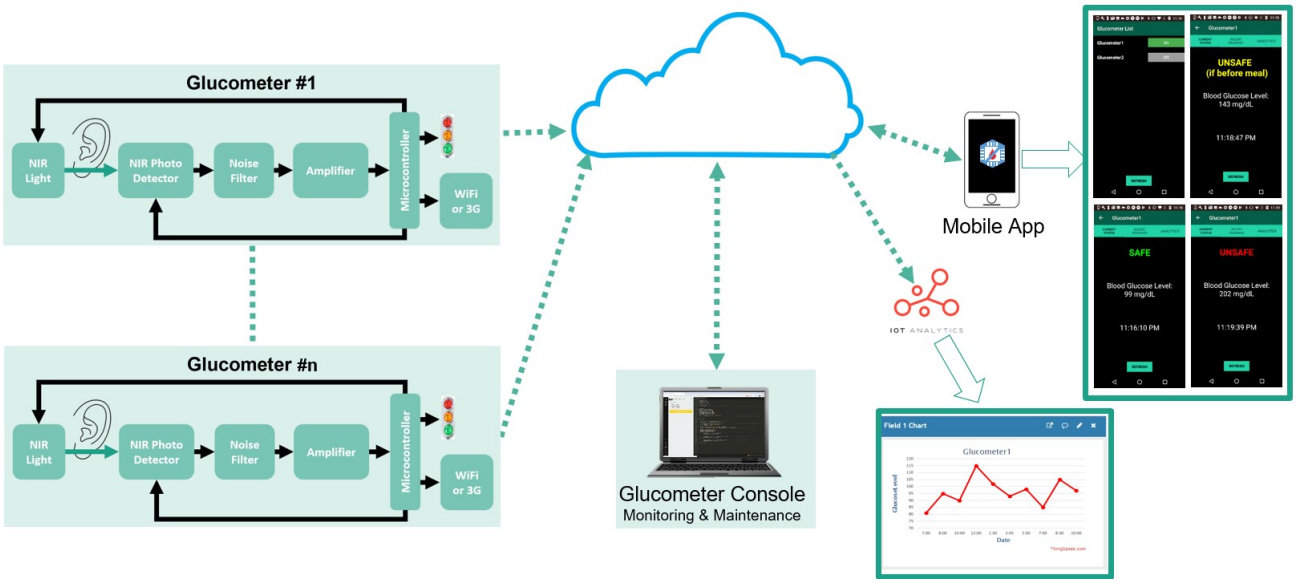


Fig. 1. Overall architecture of Ichor

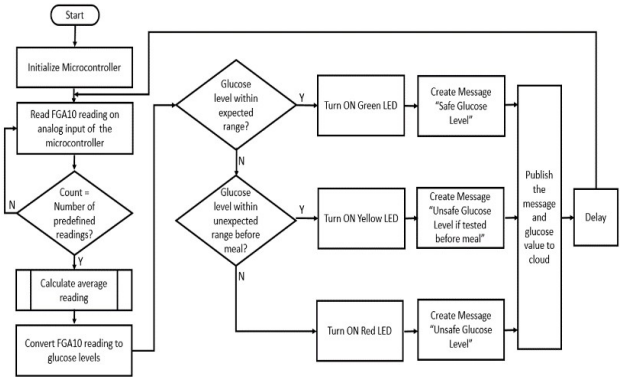


Fig. 2. Flowchart of the firmware programmed into glucometer devices

All of the components are assembled on a breadboard except the NIR LED and photodiode. After various experiments, it was found that connecting sensors to the earlobe produced more accurate results than the fingertip. Both the NIR LED and photodiode were installed on an ear clip. Fig. 3 shows the prototype of the glucometer device and the ear clip.

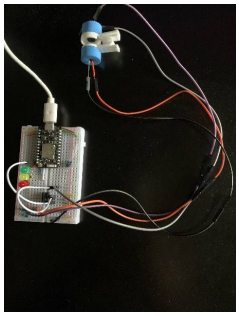


Fig. 3. Prototype of Glucometer device

The data read from the sensor is sent to the Cloud using Wi-Fi or 3G mechanisms. For the purposes of the prototype, the Particle Cloud service was used due to the simplicity of the integration; however, any other commercial cloud services can be used.

The glucometer console is used to maintain all of the glucometers. The console provides the list of all glucometer devices and their current status. Also, it supports over-the-air firmware updates.

The system supports local and remote dissemination mechanisms for alerts. The local alert mechanism uses red, yellow, and green LEDs. When glucose levels are at a safe level, the green LED on the device is turned on. When glucose levels are at an unsafe value (if before a meal), the yellow LED is turned on. Finally, the red LED is turned on when glucose levels are at unsafe levels (whether it is before or after a meal). These levels can be specifically set for each individual based on their physician's recommendation. For the prototype, the values indicated in Table 1 are used to turn on different LEDs. The remote alert dissemination mechanism involves an Android mobile app. The high-level flowchart of the mobile app is shown in Fig. 4.

TABLE I. STATUS AND LED COLOR BASED ON GLUCOSE RANGES

Status	Glucose Range (g)	LED Color
Safe	$80 \leq g \leq 130$	Green
Unsafe (If before a meal)	$130 < g \leq 180$	Yellow
Unsafe (At any time)	$g < 80$ OR $g > 180$	Red

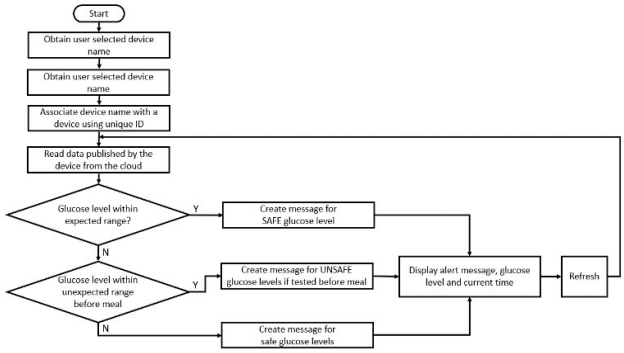


Fig. 4. Flowchart of the mobile app program

The mobile app shows the list of glucometers that the user is authorized to see. When a particular glucometer is selected, glucose level readings along with a timestamp are displayed. It also shows color-coded visual indications of safe and unsafe levels. Fig. 5 shows screenshots of the Android App.

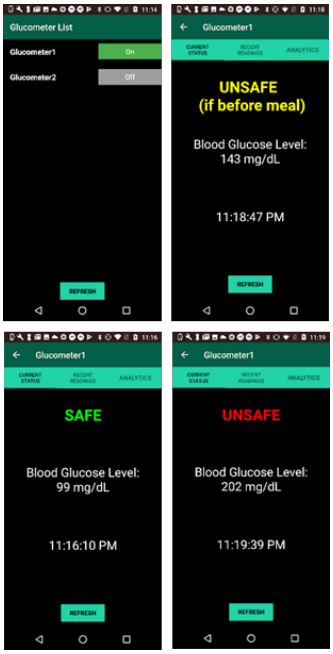


Fig. 5. Screenshots of the Android mobile App

IoT analytics, such as ThingSpeak, provides support for historical data. Fig. 6 shows a screenshot of a ThingSpeak graph.

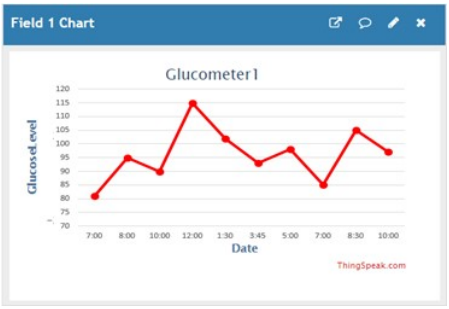


Fig. 6. IoT analytics graph from ThingSpeak

III. Mathematical Model

The Beer-Lambert law states that there is a linear relationship between the absorbance and the absorption coefficient, the optical path length, and the concentration of the solution [4] [9].

$$A = log(I_o/I_i) = \varepsilon bc \quad (1)$$

A= Absorbance
I_i = Intensity of light entering the solution
I_o = Intensity of the light output
ε = Absorption coefficient
b = Optical path length
c = Concentration

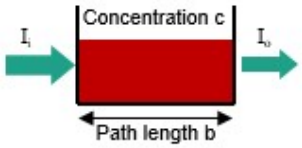


Fig. 7. Beer-Lambart Law

IV. Experiments and Data Collections

Blood glucose and dextrose monohydrate have the same chemical formula, as shown in the molecular diagrams in Fig. 8.

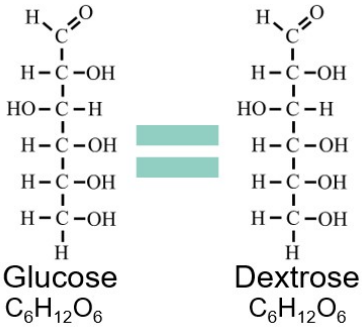


Fig. 8. Glucose and dextrose molecular diagrams

The experimental setup shown in Fig. 9 was used to collect data.

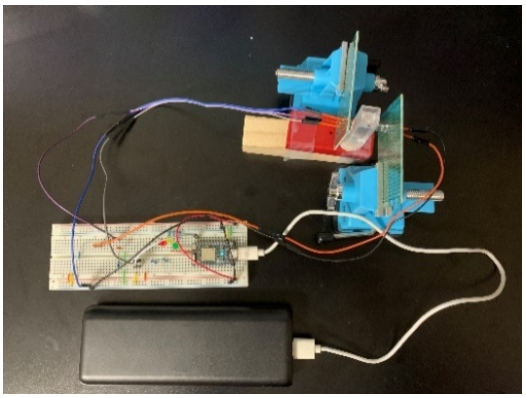


Fig. 9. Experimental setup used to collect data

Dextrose monohydrate was used to create glucose solutions of concentrations ranging from 20 to 400 milligrams per deciliter. For example, to prepare a 100 milligrams per deciliter concentration of glucose, 100 milligrams of dextrose monohydrate was dissolved in one deciliter of distilled water. 2.5 mL of each concentration was poured into a cuvette and positioned in-between the NIR light and the InGaAs detector, as shown in Fig. 9. Then, for each of the concentrations, the analog value reported in the console is collected. The analog value is divided by 4095 and multiplied by 3.3 to get the corresponding voltage value. The reason the analog value is divided by 4095 is that the analog input of the microcontroller that the author used can provide 0 to 4095 digital values. All the data collected was put into a spreadsheet, the points were graphed, and a regression analysis was performed. Fig. 10 to Fig. 13 shows four

randomly selected trails out of the dozens that were performed using the experimental setup and dextrose solution.

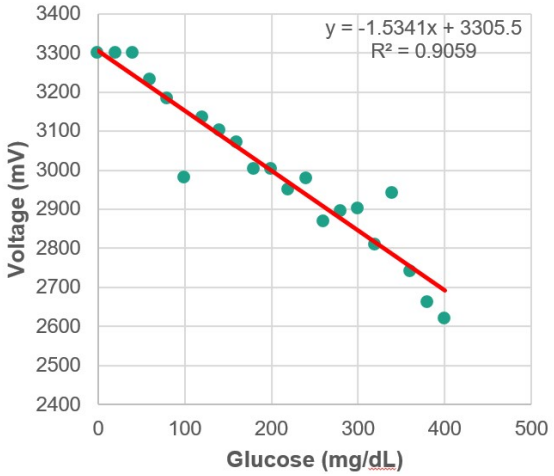


Fig. 10. Regression analysis – trial 1

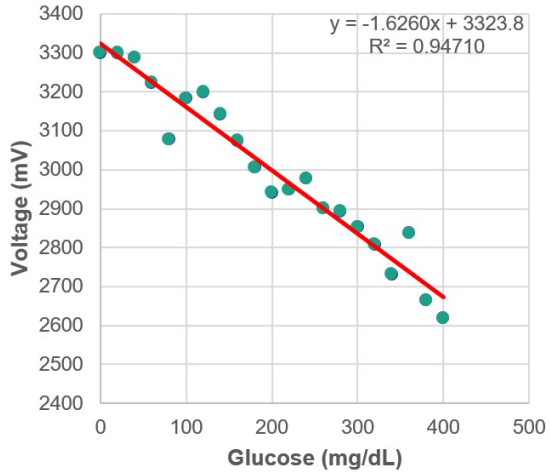


Fig. 11. Regression analysis – trial 2

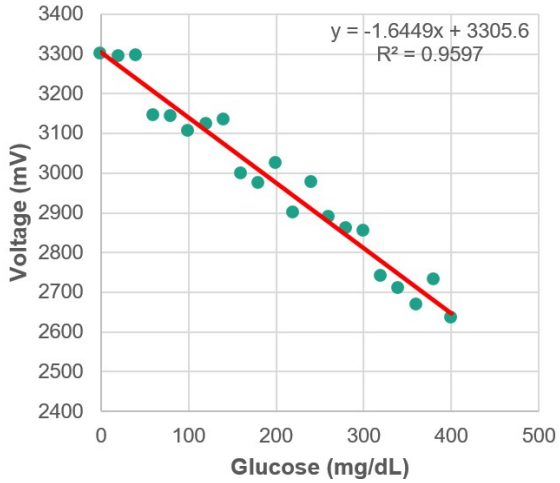


Fig. 12. Regression analysis – trial 3

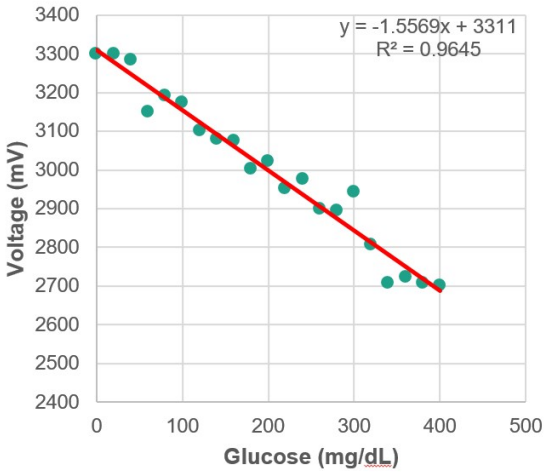


Fig. 13. Regression analysis – trial 4

The voltage values in millivolts read from the photodiode were graphed for different glucose concentrations in milligrams per deciliter. The regression analysis in Fig. 10 shows that 90.59% of the data fit the regression model. Fig. 11 shows over 94.71% of the data follow the regression. In Fig. 12 and Fig. 13, 95.97% and 96.45% of the data, respectively, follow the regression model. Overall, the randomly selected data sets indicate that the data of the measured voltage for the corresponding glucose concentration is between 90% and 96% of the regression model. This indicates that there is a strong correlation between glucose concentration and voltage.

V. Evaluation and Analysis

The Ichor device was built using a microcontroller from Particle, a NIR LED from Thorlabs (LED1550E), and an InGaAs Photodiode from Thorlabs (FGA10). An ear clip that houses NIR LED and photodiode was 3D printed. Firmware for the microcontroller was developed using the Particle Integrated Development Environment (IDE). The device sends the glucose readings to the Particle Cloud. The Android mobile app was developed to pull glucose readings from the Cloud and display glucose levels and alerts. The Ichor system is integrated into the ThingSpeak IoT analytics platform to show historical data.

Once a strong correlation between glucose concentration and voltage was established, an Android mobile app was developed to display glucose levels. The Ichor mobile app was used to further evaluate the system with a common commercial glucometer. The author collected ten random samples a day over dozens of different days of his own blood using the commercial glucometer and compared it with the data collected from Ichor. Based on this data, a bar graph was created to show the relation between the data reported by the Ichor mobile app and a commercial glucometer. Four random samples from the data that were collected are shown in Fig. 14 to Fig. 17.

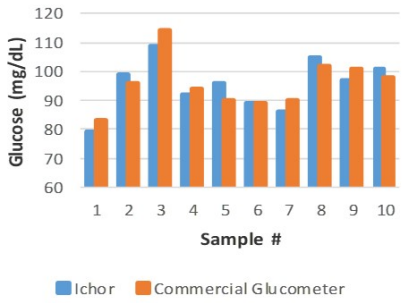


Fig. 14. Ichor versus commercial glucometer – trial 1

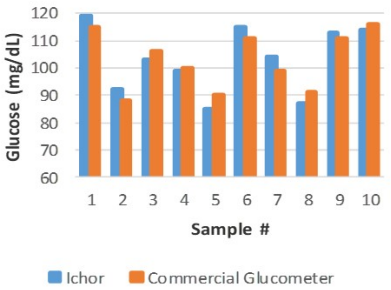


Fig. 15. Ichor versus commercial glucometer – trial 2

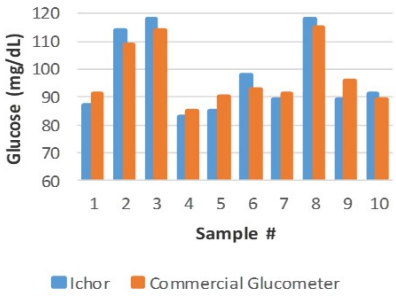


Fig. 16. Ichor versus commercial glucometer – trial 3

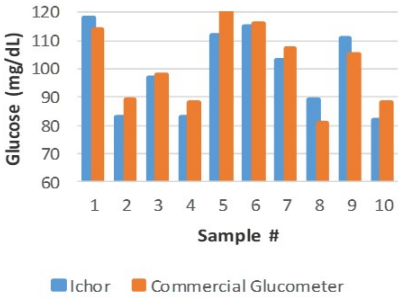


Fig. 17. Ichor versus commercial glucometer – trial 4

The error percentage of Ichor with respect to the commercial glucometer for each sample was calculated. In Fig. 14, the error percentage ranges from 0.00% to 6.67%. The error percentage of each sample in Fig. 15 is between 1.00% and 5.56%. Similarly, Fig. 16 shows a trial that has an error percentage ranging from 2.19% to 7.29%. Finally, trial 4, shown in Fig. 17, has an error percentage of 0.86% to 9.87%. Although there are outliers in the data, all four of the randomly selected iterations show that Ichor's readings closely follow those of the commercial glucometer.

After the accuracy of Ichor was proven, Ichor was integrated with ThingSpeak IoT analytics. The data reported by IoT analytics was validated by collecting glucose readings at regular intervals and comparing the data with the data reported by the analytics platform.

VI. Conclusion

Based on exhaustive testing, Ichor is a reasonably accurate and low-cost solution. Although further testing is required with a larger sample size, results show a strong relationship between actual glucose levels and data collected using the noninvasive glucose monitoring system. It costs less than $100 to build a prototype. Sending alerts is free with Wi-Fi and less than three dollars per device per month using 3G. It integrates with Cloud services enabling the development of mobile apps to send alerts to caregivers. It also integrates with IoT Cloud analytics platforms that show real-time and historical data. Alerting mechanisms built into the device will alert caregivers when glucose levels reach unsafe levels. This system can also be used in Artificial Pancreas Device Systems. Ichor will allow people with diabetes to manage their sugar levels easily and avoid hypoglycemia and hyperglycemia.

References

[1] Who.int. 2020. Diabetes. [online] Available at: <https://www.who.int/news-room/fact-sheets/detail/diabetes> [Accessed 13 August 2020].

[2] Cdc.gov. 2020. [online] Available at: <https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf> [Accessed 13 August 2020].

[3] WebMD. 2020. Hyperglycemia And Diabetes. [online] Available at: <https://www.webmd.com/diabetes/guide/diabetes-hyperglycemia> [Accessed 13 August 2020].

[4] Chemguide.co.uk. 2020. Absorption Spectra - The Beer-Lambert Law. [online] Available at: <https://www.chemguide.co.uk/analysis/uvvisible/beerlambert.html> [Accessed 13 August 2020].

[5] 2012. [ebook] Available at: <[5] https://www.thorlabs.com/drawings/41a953f64fa1c42b-F3C4B1AF-FF82-1829-E550B3E654C44020/LED1550E-SpecSheet.pdf> [Accessed 13 August 2020].

[6] 2012. [ebook] Available at: <[6] https://www.thorlabs.com/drawings/41a953f64fa1c42b-F3C4B1AF-FF82-1829-E550B3E654C44020/FGA10-SpecSheet.pdf> [Accessed 13 August 2020].

[7] Docs.particle.io. 2020. Electron Datasheet | Datasheets. [online] Available at: <https://docs.particle.io/datasheets/electron/electron-datasheet/> [Accessed 13 August 2020].

[8] R. Hotmartua, P. W. Pangestu, H. Zakaria and Y. S. Irawan, "Noninvasive blood glucose detection using near infrared sensor," 2015 International Conference on Electrical Engineering and Informatics (ICEEI), Denpasar, 2015, pp. 687-692, doi: 10.1109/ICEEI.2015.7352586.

[9] D. F. Swinehart, "The Beer-Lambert Law", Journal of Chemical Education, vol. 39, pp. 333, Jul 1962.

[10] K. T. Sithara Surendran and T. Sasikala, "Sensor system based Non-Invasive System to Measure Glucose Level in Human Body," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2019, pp. 203-207, doi: 10.1109/ICCMC.2019.8819766.

[11] C. Pasquini, "Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications," *J. Braz. Chem. Soc*., vol. 14, no. 2, pp. 198–219, 2003.

New Routing Protocol for Half-Duplex Cognitive Radio Ad-Hoc Networks over IoT Environment

Khalid A. Darabkh[1] and Oswa M. Amro[2]
[1]Dept. of Computer Engineering, The University of Jordan, Amman, 11942, Jordan, Email: k.darabkeh@ju.edu.jo
[2] Department of Electrical Engineering, The University of Jordan, 11942, Amman, Jordan, Email: oswa.amro@hotmail.com

***Abstract*—In this paper, we consider non-time slotted cognitive radio ad-hoc networks in order to propose a routing protocol without a common control channel, thereby adhering to practicality. This is performed using multicast and unicast transmission of the control packets. The performance of this protocol is studied and compared with a quite relevant and recent protocol, called Probabilistic and Deterministic Path Selection (PDPS), using a special simulator, which is built in Java language and developed upon PDPS's simulator. The PDPS protocol creates two paths between the source and destination while using all the channels in the network. Interestingly, their results are promising in terms of throughput. On the other hand, our protocol builds as many paths as possible between the source and destination. Besides, in our protocol, we use all the channels in the network. The performance metric considered is the throughput whereas our results are even better than those of PDPS protocol.**

Keywords— Cognitive Radio Ad-Hoc Networks (CRAHN), no Common Control Channel (CCC), non-time slotted CRAHNs, and CRNs' Routing Protocol.

I. INTRODUCTION

Internet of Things (IoT) is defined as the connection of diverse items over the internet, operating with various communication technologies [1]. From which, there are wired and wireless technologies, the latter is more flexible, thus, more popular [1]. Nevertheless, major concerns arise in this case, for example there is the existence of massive data, and the shortage in spectrum availability. Interestingly, researchers have a desire to solve these challenges using cognitive radio networks (CRNs) [2]. CRNs search the spectrum for available bands, then dynamically change the transmitting parameters of the devices to opportunistically access the spectrum [1], [3]. There are many potential applications for CR-based IoT devices, such as healthcare applications, social activities (e.g., smart traffic lights), environmental related applications, in-home applications, smart grids, smart cities, and internet of vehicles [1].

Half-Duplex multi-channel non-time-slotted CR Ad-Hoc Networks (HD-CRAHNs) are considered a practical spectrum-efficient communication model. In this model primary and secondary networks are asynchronous. Specifically, Ad-hoc networks consist of wireless nodes communicating via multi-hop paths, without an infrastructure. Moreover, cognitive radio technology allows the nodes in the network to use the empty frequency bands without having bands of their own. In CRAHNs, a source node communicates with a destination by building a multi-hop path between the two nodes. All the nodes can communicate without a fusion center. This guarantees high coverage of the network, low cost, and easy maintenance.

At this point, it is apparent that having CR-based devices in IoT applications is very promising. Thus, we were motivated to work on an Ad-Hoc CRN that can be used in IoT applications, such as smart cities. In our work, we propose a routing protocol for multi-channel non-time slotted CRAHNs.

To create a routing protocol that is aware of the channel's availability and the Primary Users' (PUs) activity (i.e. spectrum-aware), one should think of the sensing and the spectrum allocation processes along with the routing process. This means creating cross-layer routing protocols is essential to enhance the performance of CRNs, especially for meeting the demands of multi-media applications [4]. The text below highlights two of the most relevant cross-layer routing protocols for CRNs.

A cross-layer routing protocol named CLRP is proposed in [5], which considered all available channels at every SU node. This protocol formed a full path before deciding which channels should be sensed by the PHY-layer of each node. Similarly, a protocol named Probabilistic and Deterministic Path Selection (PDPS), that extremely enhances the performance CLRP, is proposed in [6]. In which, the authors built two paths between the source and the destination and they let the source decide which path has better throughput.

The proposed protocol introduces new concepts and make modifications to CLRP. It also considers the existing CLRP modifications, such as PDPS. We name our protocol Multi-Cast-based Half Duplex Routing Protocol (MC-HDRP). Remarkably, for page length constraints and because PDPS outperforms CLRP, we compare our work only with PDPS.

The rest of the article is organized as follows: Section II illustrates the proposed protocol. The system and protocol's assumptions are detailed in Section III and VI, respectively. Mathematical model is presented in Section V. Extensive details of the proposed protocol are provided in Section VI. Section VII presents the simulation results. The article is concluded in Section VIII.

II. THE PROPOSED PROTOCOL

Interestingly, both CLRP and PDPS consider the probabilistically and the deterministically available channels. Where the probabilistically available channels have specific idle probability, assumed to be found after conducting long term monitoring processes. However, the deterministically available channels represent a small set of channels that is assigned to each node based on the geographical area and is sensed periodically by the node. This is in order to mimic the process of serving other paths than the one in question and gaining the advantage of having channels ready to use whenever needed. In our protocol we do not assign this set of channels to every node, rather we consider all the channels only before the transmission starts. As we think that sensing this set of channels periodically wastes a portion of the slot time and negatively affects the transmission period, i.e., it reduces the throughput. Besides, we consider other paths to be served in the initial load inside each time slot. Hence, only the used channel is going to be sensed periodically using Listen Before Talk (LBT) protocol. Moreover, we study the

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

probabilistically available channels too, in order to represent the activity of other Secondary Users (SUs) in the network. SUs are assumed to be arriving according to a Poisson distribution. This is deferent from CLRP and PDPS where they used the probabilistically available channels to mimic the behavior of PUs. Remarkably, to be fair, we removed the assignment of deterministic channels to the nodes in PDPS that we compared our work with.

We add timing to our protocol and simulate the existence of PUs. We use a delay variable and a set of files with PUs existences. However, for testing purposes, we represent the activity of PUs as various periods of time taken randomly from every channel. Besides, each channel can be used by various PUs, each one at a time

III. System Model

Each selected path in our protocol consists of a source node, intermediate SU nodes, and a destination node. Only one channel is used between each two nodes, i.e., each SU node has only one transceiver. Specifically, the used channel in each hop is chosen from all the channels in the network after the sensing process is conducted. Given that the availability of each channel has a specific probability, which is called idle probability. In our study, there exist three types of channels: "Available" (idle probability = 1) such that there is no PU nor SU, "Not available" (idle probability = 0) which only means a PU exists, and "Probabilistically available" (0 < idle probability < 1) such that SUs exist in the channel with the given idle probability. The last type of channels was named "unknown to be available" in [5], [6] but for PUs activities not SUs. Interestingly, the SU nodes should be contending, therefore, SUs activity will only be considered in the "probabilistically available" channels not in any other type. Obviously, the state of each channel is found after the first sensing process is conducted.

The first sensing process is conducted to all the channels before the network settles down. Then, there is a sensing process for the "probabilistically available" and the "available" channels at each node's position before message transmission. This is followed by a periodic sensing process for the channel used in transmission. Additionally, if the channel gets busy while the SU is transmitting (i.e., a collision) the transmitted packet will be dropped, and it will not be transmitted to the neighbor. This is because we do not have acknowledgments. Therefore, the node will look for another channel to transmit its packets over, only after the sensing period comes (using the LBT protocol). This is how we avoided the use of a CCC.

IV. Protocol's Assumptions

- We assume that there exist various PUs in the network, each one can occupy any channel. Every channel can be occupied by one user at a time.
- Given that a PU can tolerate interference for a specific interval, called the Tolerable Interference Delay (TID). We assume that TID is 1 second long. An SU is preferred to stop transmitting and monitor the PU activity every TID. This is done using LBT protocol for the sensing /transmitting tradeoff with overlay dynamic spectrum access model.
- We use energy-based and waveform-based detection. The first technique is used to detect whether there is a user or not. Only if there is a user, the second technique would be used; in order to check if that user is an SU. If it was another SU, then the SU in question will try to access the channel anyway but with throughput multiplied by the idle probability. Moreover, if the given SU was already transmitting on the channel and another SU appears, it will keep occupying the channel until the next sensing period in HD mode. This idea is inspired by [7], [8], where the authors handled many problems facing CRAHNs like SU's frequent handoffs.
- We simulate the activity of SUs only as if they are arriving according to a Poisson distribution with known birth and death rates. Intriguingly, we simulate the PUs as actual nodes occupying the channels randomly.
- We assume that all the channels have the same bandwidth (BW). Additionally, we adopt a bandwidth model that has 1 bit per Hz per second if the received Signal-to-Interference-plus-Noise-Ratio is higher than a predefined threshold, i.e., $R = BW\ [bit/sec]$.
- No Common Control Channel (CCC) is required. This is an advantage, because we avoid reserving a channel for sending control packets (i.e., spectrum efficiency). Besides, this way we follow the diffenetion of a CRN, as it is designed to operate over the empty channels.
- Some requirements are assumed to be known, such as the location of our SU nodes, the set of all channels in the CRN, the expected available time of the channel and the birth and death rates of contending SUs. Besides, the keep-out distance of channels which indicates the locations of contending SUs is also known. Note that, in PDPS, the information that should be known is about PUs' network. Thus, we have a more practical system because the cooperation between SUs and PUs is not required.
- We have the maximum power constraint of idle and probabilistically available channels. In this constraint, the channel is considered available only when its required power is less than the node's maximum transmission power.

V. Mathematical Model

In our protocol we adopt equations used by [5], [6], thus there will be no need to re-write the equations in this work. However, the most important equations are mentioned, such as the transmsission time, which is given by

$$T_{T_x max} = TID - L_w - ST_{Ch_1} - ST_{Ch_2} - \alpha\left|f_1 - f_2\right| \qquad (1)$$

where (TID) is the periodic time slot, it is actually the PU's tolerable interference delay and it is assumed to be 1 sec. The initial load is (L_w) which is the time to serve other paths. The sensing time is (ST). The frequency step or switching time constant (α) is in seconds per Mega Hertz. Finally,(f_1) and (f_2) are the central frequencies of the two channels (Ch_1) and (Ch_2), respectevly.

To get the throughput in bits per second we multiply the transmission time ($T_{Tx}max$) by the bandwidth of the channel (BW) then divide it by the number of time shares, such as the following equation.

$$TH_{max} = \frac{T_{T_x max} \times BW}{\text{time shares}} \qquad (2)$$

A time share means the time slot where we have simultaneously active nodes in a path. But another definition

states that it is the existence of a split in time between the activities of the nodes in the path. Based on both definitions, in HD we have the number of time shares equals to the number of hops in the path.

VI. Details of the MC-HDRP Protocol

A. Route Request Packets in MC-HDRP

A setteling time was added to every node to guarantee that the node senses all the channels and knows its neighbors.This delay is added before the start of any transmission.

At the beginning of any transmission, the source senses a channel from its surroundings to detect the PUs and the secondary neighbors. The detection relies on the transmission of PUs and on the leaked oscillator signals that accompany the sensing process of SUs [7], [9], [10]. Hence, each leaked signal must have a unique ID. This idea is an optimistic thinking in the light of [11]–[14]. The oscilator power leakage of any reciever (or sensing device) is discussed in [15] and [16].

The source sends a multicast Route Request Packet (RRQP) over the channels that have SUs sensing them. The transmission happens without repetition, i.e., the neighbors who received the RRQP will be eliminated from other transmissions throughout the channels. This can happen because the SUs can be detected based on an ID of their signals, thus, multicast or unicast is used. The idea of multi-casting was mentioned in [17]. Remarkably, an SU neighbor can be detected over various channels because the nodes sense the channels and transmit over them in diferent orders.

The source adds its neighbors to a queue and each neighbor will add its neighbors to the queue without repetition. Most importantly, this queue is sent in the RRQP. This queue is devided into “next in queue” and “transmitters” lists as seen in Table I and as will be explained in the following scenario.

TABLE I. The RRQP Contents for Multi-Casting in MC-HDRP and MC-HDRP-CA.

Channel ID	**Node ID**	**Upstream channel**	**Upstream node**	**Downstream quality**	**Entity length**
The downstream channel	This node ID	Previous channel	Previous node	The time for transmitting from the slot	In bits
. .	. .	. .	. .	. .	. .
The Queue: …					
The transmitters’ list: …					

1) Scenario for RRQP delivery

Figure 1 shows a scenario for RRQP delivery in our proposed protocol. In particualer, a queue and a transmitters’ list are two entries added in the RRQP along with an entry for every downstream channel to carry the channel’s downstream quality, upstream node ID, corresponding upstream channel ID, entry’s length (indication of the message length), and this node’s ID.

Its noteworthy to mention, before explaining the example, that the neighbors were found in the settling time, however, before transmitting the RRQP the node must sense every channel to recognize the existing neighbors in that channel. Therefore, the queue depends on which node was discovered first and on which channel. Additionally, only in the example, the nodes are discovered from the west of the node and clockwise. Besides, we assume in the example that each node is found on a separate channel at every transmission process.

The following text explains the example given in Figure 1.

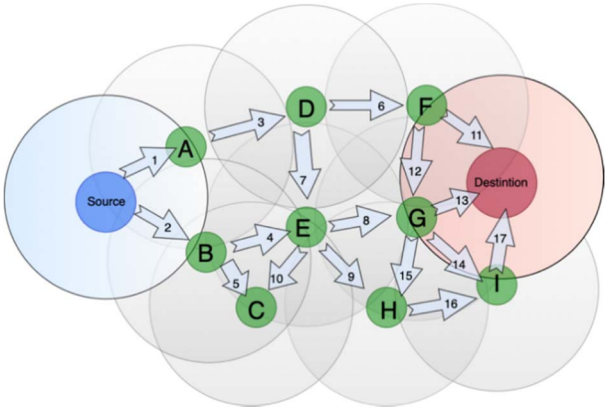


Figure 1. RRQP Delivery in the Route Discovery Stage

- In the settling time, the source (1) discovers nodes “A and B” each on a separate channel. Then, (2) the source senses the first channel and adds node “A” to the queue, right before (3) transmitting over the first channel. This is followed by (4) sensing the next channel, finding node “B”, then adding it to the queue (after “A”). At this point, node “B” will receive the RRQP with the new queue over the second channel. the transmitters’ list has only the source in it.
- The first node in the queue has the priority to send its version of the RRQP to its pre-defined neighbors. In this case the first node is “A” and (recall) it sees itself as the only one in the queue. Consequently, node “A” has “B and D” as neighbors known right before transmission starts (i.e., node “A” will sense the channels separately to find the neighbors). Then, (assuming the west-clockwise discovery in the example) “A” will send the RRQP to node “D” then node “B” because they are not in the transmitters’ list received from the source. Additionally, node “A” will be moved from the queue to the transmitters.
- Node “B” will keep waiting until it receives the new RRQP from node “A” (the first in queue). In the same manner node “B” discovers that it has “A, E and C” as neighbors on separate channels. Node “A” is in the transmitters’ list (from the updated RRQP). Therefore, node B will send the RRQP only to nodes “E” then “C” each with a queue, (D and E) then (D, E and C), respectively. It is important to note that node “B” will consider both RRQPs received from the “source” and node “A”. As for the quality, “B” will get the best quality for every downstream channel. However, for the queue and the transmitters’ list node “B" will get the latest received ones (from A). Interestingly, the transmitters list at “B” has: “Source, A and B”.

This process keeps going in the same manner. Thus, for space constraints we will jump to node “I” without considering the details.

- Node “I” handles the updates of the transmitters and the queue lists received from node “H” not “G” because “H”

is the latest to transmit. Node "I" has "H, G, and Destination" as neighbors. Thus, it will send only to the destination; because "H" and "G" are in the transmitters list. The queue has no entries and the transmitters' list has: "Source, A, B, D, E, C, F, G, H and I".

B. Route reply packets in MC-HDRP

The destination processes the received RRQPs and creates its Route Reply Packet (RREP) with the best received RRQP for every channel. Then it will send the RREP (one message item per node) to its neighbors. Only the ones who sent the RRQP to the destination. As for transmission, the "destination" uses the best channels in terms of throughput. This is done to create multiple paths.

The reverse paths have consequent nodes working simultaneously (i.e., the paths are multi threaded). Additionally, each node in the reverse path will take the best channel to transmit the RREP to the upstream node. The throughput of the path is the minimum of all the hops in that path (i.e., bottleneck throughput). At the end of path discovery stage, the source decides which path has the maximum throughput.

VII. Simulation Results

This section begins by introducing the simulation parameters and the performance metric used in the comparisons between protocols.

A. Simulation Environment

The simulations were conducted using a special simulator built in Java language that was heavily edited over CLRP and PDPS's simulators. The computer used was Apple MacBook with 1.8 GHz Intel Core i5 processor and 8 GB 1600 MHz DDR3 RAM.

The simulator was built with various classes, such as network, node, channel, basic message, basic message item, control message, message item, data message, packet, path, and primary user. The thread used in the simulator was the RREP delivery thread.

B. Simulation Parameters

The default values of some parameters are listed in Table II, however, any change in the parameters is mentioned before every figure. Each point in the figures is an average of 100 trials. The random values follow uniform distributions except the arrival of SUs, which is Poisson. However, in PDPS the arrival of PUs is considered.

C. Performance Metric

We use the throughput as our performance metric. The throughput in our study can be described as the transmission time in each time slot multiplied by the channel bandwidth and divided by number of time shares in the path. We can express the throughput in terms of bit/sec/Hz if we did not multiply by the bandwidth; as we have the same channel bandwidth in our simulations. More to the point, the theoretical description of the throughput is the average rate of successfully delivered packets per second.

D. Discovery Phase

This section provides performance evaluation of the proposed protocol in the route discovery phase.

Figures 2-a, 2-b and 2-c show the effect of increasing the number of channels on the behavior of our proposed protocol compared with PDPS in low idle probability conditions, moderate and high, respectively. Both protocols have fixed throughput after the number of channels is enough; because the extra channels will not be used to increase the throughput.

TABLE II. TABLE 1. DEFAULT SIMULATION PARAMETERS.

Parameter	Value
Number of SU Nodes	50 nodes
PU TID	1 second
Number of Channels	12 channels
Transmission Range of SUs	50m
Transmission Range of Contending SUs, or PUs in PDPS	500m
SUs Network Dimensions	200m×200m
Contending SUs Network Dimensions, or PUs Network in PDPS	1000m×1000m
Idle Probability $P_r(H_0)$	As given in each graph (low 0.1, moderate 0.5, and high 0.9)
SU source position	At (0, 0)
SU destination position	At (200, 200)
Number of PU nodes	5 nodes
Initial load [5]	Random between 0.1 and 0.7 seconds
Switching Constant α.	1 milli sec / MHz
Sensing time of a channel [5]	Random between 1 milli sec and 100 milli sec Max 1/10 of the TID for the channel
Channel Bandwidth	6 MHz

MC-HDRP has higher throughput compared to PDPS, with increment ratio equal to 2.84% (with idle probability = 0.1). This can be explained as MC-HDRP uses all the possible paths between the source and the destination. Thus, the chances of gaining better throughput are higher.

As for the second experiment, default values are used and number of SU nodes is variable. In Figures 3-a, 3-b and 3-c, as the number of nodes increases then the connectivity increases, which results an increment in the throughput in both protocols. The throughput will not increase significantly, because of the interference effect on the throughput when number of SUs increase.

VIII. Conclusions

The use of CR technology in Ad-Hoc networks guarantees using the spectrum efficiently to handle the increasing demands on wireless communications, especially IoT applications. In this paper, a practical HD-based CRN routing protocol is introduced. The proposed protocol does not use a CCC, rather it uses multi- and uni-cast for sending the RRQPs. Our protocol (MC-HDRP) outperforms the PDPS protocol introduced by [6] in terms of throughput. This is because our protocol uses all the possible paths and avoids having frequent handoffs by the SUs.

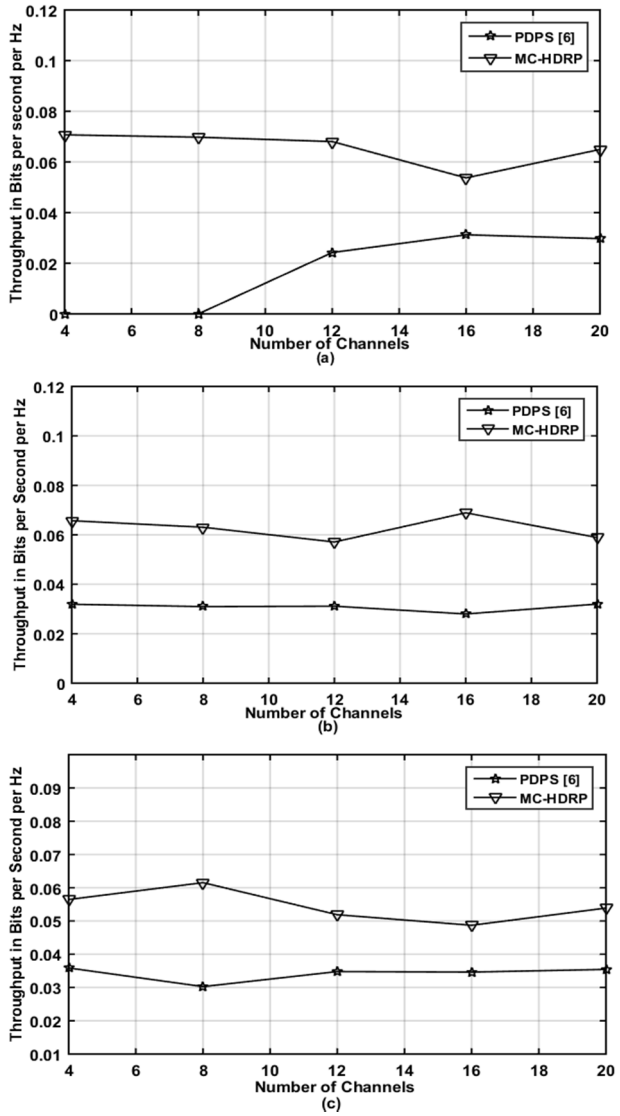


Figure 2 Discovery Throughput vs Number of Channels: (a) Idle Probability 0.1, (b) Idle Probability 0.5, and (c) Idle Probability 0.9

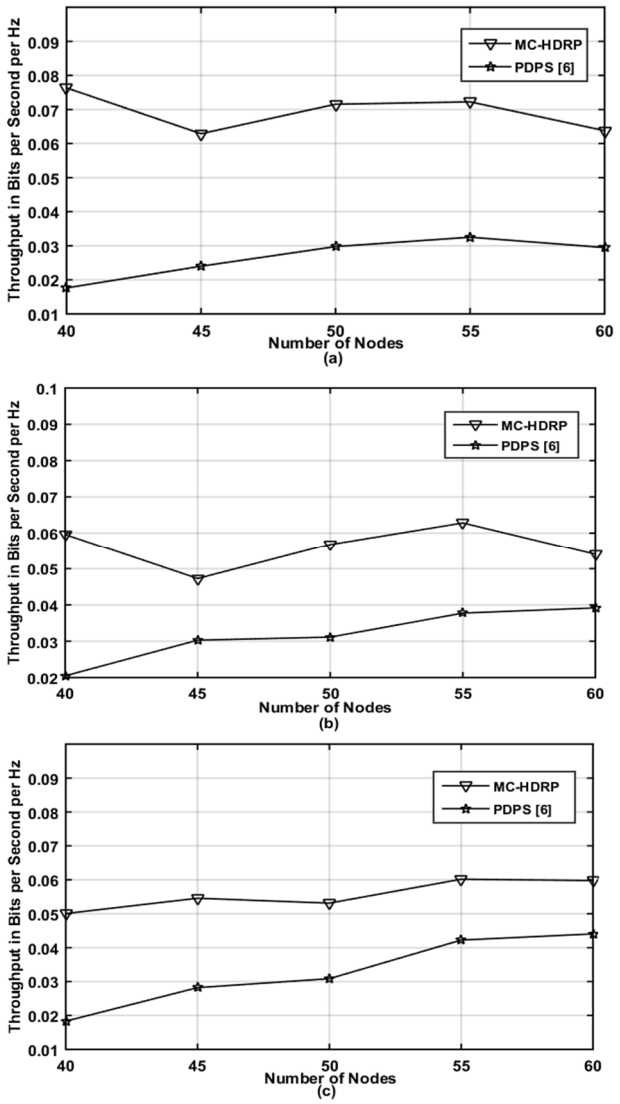


Figure 3 Discovery Throughput vs Number of Nodes: (a) Idle Probability 0.1, (b) Idle Probability 0.5, and (c) Idle Probability 0.9

REFERENCES

[1] Wafa'a Kassab and Khalid A. Darabkh, "A-Z Survey of Internet of Things: Architectures, Protocols, Applications, Recent Advances, Future Directions and Recommendations," *Journal of Network and Computer Applications*, vol. 163, p.102663, August 2020.

[2] Haythem Bany Salameh *et al.*, "Channel Assignment Mechanism for Cognitive Radio Network with Rate Adaptation and Guard Band Awareness: Batching Perspective," *Wireless Networks*, vol. 26, no. 6, pp. 4477-4489, 2020.

[3] K. A. Darabkh, O. M. Amro, H. Bany Salameh, and R. T. Al-Zubi, "A-Z Overview of the In-Band Full-Duplex Cognitive Radio Networks," *Comput. Commun.*, vol. 145, no. September 2019, pp. 66–95, 2019.

[4] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, 2007.

[5] R. Saifan, A. E. Kamal, and Y. Guan, "A cross-layer routing protocol (CLRP) for cognitive radio network," in *GLOBECOM - IEEE Global Telecommunications Conference*, 2013.

[6] R. Saifan, A. M. Msaeed, and K. A. Darabkh, "Probabilistic and deterministic path selection in cognitive radio network," *IET Communications*, vol. 13, no. 17, pp. 2767–2777, November 2019.

[7] W. Cheng, X. Zhang, and H. Zhang, "Pilot-based full-duplex spectrum-sensing and multichannel-MAC over non-time-slotted cognitive radio networks," in *Proceedings - IEEE INFOCOM*, 2017.

[8] W. Cheng, W. Zhang, L. Liang, and H. Zhang, "Full-duplex for multi-channel cognitive radio ad hoc networks," *IEEE Netw.*, no. 99, pp. 1–7, 2018.

[9] A. Chaman, J. Wang, J. Sun, H. Hassanieh, and R. R. Choudhury, "Ghostbuster: Detecting the presence of hidden eavesdroppers," in *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM*, 2018.

[10] C. Shao, W. Jang, H. Park, J. Sung, Y. Jung, and W. Lee, "Phantom Eavesdropping with Whitened RF Leakage," *IEEE Wirel. Commun. Lett.*, 2019.

[11] Y. Li, X. Chen, J. Chang, and Y. Lin, "The Physical Layer Identification of Communication Devices Based on RF-DNA," in *Advanced Hybrid Information Processing*, 2019, pp. 452–460.

[12] B. D. Stone, "Comparison of radio frequency distinct native attribute and matched filtering techniques for device discrimination and operation identification," 2016.

[13] T. McGrath, I. E. Bagci, Z. M. Wang, U. Roedig, and R. J. Young, "A PUF taxonomy," *Appl. Phys. Rev.*, vol. 6, no. 1, p. 11303, 2019.

[14] B. Chatterjee, D. Das, S. Maity, and S. Sen, "RF-PUF: Enhancing IoT security through authentication of wireless nodes using in-situ machine learning," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 388–398, 2018.

[15] G. Zhao, W. Zhang, and S. Li, *Advanced sensing techniques for cognitive radio*. Springer, 2017.

[16] B. Wild and K. Ramchandran, "Detecting primary receivers for cognitive radio applications," in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005.*, 2005, pp. 124–130.

[17] N. Avinash and S. Tamilarasan, "Performance Analysis of Different Routing Protocols in Cognitive Radio Networks," *Perspect. Commun. Embed. Signal-processing-PiCES*, vol. 1, no. 10, pp. 164–168, 2018.

Low-Cost Contact Thermometry for Screening and Monitoring During the COVID-19 Pandemic

Narasimha Sai Yamanoor
Kenmore, NY, USA
yamanoorsai@gmail.com

Srihari Yamanoor
San Jose, CA, USA
yamanoor@stanfordalumni.org

***Abstract*—A key component of the UN Sustainable Development Goals (SDG) is goal 3, Good Health, and Well-Being. Fundamental to the accomplishment of this goal is women's health. Pregnant women and women raising infants would benefit from early screening. In socially disadvantaged areas, patients may not have ready and frequent access to formal healthcare or screening solutions. In such instances, simple solutions that allow for self-monitoring can help. Additionally, when provided with devices that promote positive behavior modification such as sensor-enabled wearable devices, other advantages may accrue. Data collection from multiple subjects for screening and contact-tracing may have potential use in the ongoing COVID-19 pandemic and beyond. A low-cost, contact thermometer solution based on a silicon bandgap temperature sensor that allows for personal screening is described using a Proof-of-Concept solution.**

Keywords—UN Sustainable Development Goals, Temperature Measurement, Wearables, Health Trends, Pregnancy, Neonatal Care, Contact Thermometer

I. INTRODUCTION

Within the UN Sustainable Development Goals (SDG), Goal 3, Good Health and Well-Being, several areas of human health have been recognized as having been part of progress, but still requiring acceleration [1]. Maternal Health and Child Health are among the stated areas [2]. Research indicates that maternal health and child health may be linked, with maternal health issues, consequently affecting fetal and child health [3] [4] [5]. The urgency for improvements in maternal health and reduction in mortality has been demonstrated [6]. While the specific effects of COVID-19 on pregnant women are not well understood currently, cautious approaches are recommended in preventive and post-diagnostic care [7].

The focus of the current work is in the demonstration of the feasibility of a low-cost open-source solution for self-care and monitoring for a specific cohort. The solution may be expanded to, include other cohorts and additionally allow for the design of population health intervention using machine learning techniques.

Medically underserved populations [8] would be better served by the availability of innovative, cost-effective diagnostic and monitoring devices [9] with at least an acceptable level of performance enhanced by technological advances in sensing. Maternal temperature measurement and monitoring have been studied and reported in clinical studies [10]. Temperature, a basic vital sign recommended for monitoring during pregnancy [11] is also indicated for pregnant women diagnosed with COVID-19 infection [7].

II. BACKGROUND

A. Project Objectives

The team had the following key objectives while considering the project planning and execution:

1) The solution should serve an underrepresented demographic that requires customizable, cost-effective solutions.
2) In keeping with the general principles of the Maker Movement [12], the design should be open, and open hardware and software should be used whenever possible. Components should be readily available to allow for mass production and distribution.
3) For processing and testing capabilities, an affordable Single Board Computer (SBC) or microcontroller, or preferably, an open-architecture device will be used.
4) The devices should render the highest acceptable screening performance at the lowest optimal cost for widespread adoption.
5) The design should be flexible to be adapted for different implementation styles, such as a handheld versus a wall-mounted device, or a wearable device, and allow for use under varying social distancing protocols.
6) Proof-of-Concept (PoC) solutions that are proposed, demonstrated, and published will aid any team worldwide to seek funding, and make efforts to realize the project's completion, maximizing the chances for success and implementation.
7) The primary objective for the project to be a response to the COVID-19 pandemic, but the solution should be extensible to non-pandemic situations as well.
8) The solution should serve individual end-users, but also lend itself to artificial intelligence techniques for pattern recognition in larger

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

populations for incidence and contact tracing, as well as other trends analyses.

9) The primary objective of the device is to function as a screening device, and not as a clinical device. Regulatory approvals and commercialization are currently outside the purview of this work.

B. Temperature Measurement

Body Temperature Measurement has been identified among the vital monitoring and diagnosis parameters for COVID-19. Fever is a key symptom, which, when it manifests, appears 2 – 14 days after infection [13]. Body Temperature Measurement is the basis for the design of a phone application along with analysis and dashboarding tools that use AI to identify infection trends [14].

Fever monitoring recommendations vary largely based on national, regional, and local governments. Notably, frequent temperature measurement is a component of most such recommendations and guidelines [15], [16], [17], [18].

C. Target Population

Non-contact and contact thermometry solutions work on different sets of principles and in the COVID-19 pandemic, both have distinct applications. For instance, non-contact thermometers are used to abide by workplace screening guidelines for which the team has demonstrated a distinct, proof-of-concept solution through [19], guidelines are available for temperature screening for pregnant women diagnosed with COVID-19 [7] and for screening during maternity in general [11], for which contact thermometry might offer pertinent solutions.

D. Types of Temperature Measurement

Body Temperature Measurement can be through non-contact or contact techniques.

1) The two non-contact thermometry device types are

a) Thermal Imaging Systems

b) Infrared Thermometers [20].

2) Contact thermometers use the conduction mode of heat transfer. They are of four types:

a) Thermocouple

b) Resistance Temperature Detector

c) Thermistor [21]

d) Silicon Bandgap Temperature Detector [22]

The various types of thermometers have applications for specific measurement necessities. Their application in human body temperature are discussed in scientific literature [23], [24], [25], [26]. For this paper, in accordance to maximize performance while minimizing cost, the Max30205 [27] was selected for the PoC phase.

E. Non-Contact Vs. Contact Temperature Measurement

Both non-contact and contact body temperature measurement techniques have their applications. To ensure social distancing during the COVID-19 pandemic, non-contact thermometers have been widely recommended [20]. The team has previously designed a PoC solution for low-cost non-contact thermometry [19]. Contact thermometers similarly have specific applications, both in response to COVID-19 and beyond. The key differences are as follows:

1. Non-contact thermometers measure surface temperature, while contact-thermometers may measure internal body temperature, based on the location of measurement [28]. The Max30205 measures surface temperature through contact.
2. When screening multiple humans, such as in retail establishments, schools, factories, and similar locations where human density and interaction is high, non-contact thermometers are preferable to avoid undesirable contact or sharing of thermometers.
3. For pregnant women, the effects of COVID-19 are poorly understood. Having a personal, wearable, or portable contact thermometer will allow pregnant women, and other potential at-risk cohorts to self-screen through non-invasive, skin surface temperature measurements. Non-contact thermometers measure skin surface temperature and accounting for factors affecting contact thermometer-based skin surface temperature measurements [30], for screening purposes, the two modalities are comparable.
4. The contact-thermometer can be continued to use post-pregnancy to monitor the health of the mother and the infant [31]. Skin temperature measurement for infants can be an indicative and screening tool in infant care.
5. Personal thermometers, when not shared with other adults, and only used on self or infants after childbirth can minimize the risk of infection spread.
6. When data collection and sharing options are provided with the device and shared voluntarily by the subjects, it can add to the body of knowledge regarding screening with skin temperature during pregnancy and after childbirth and can be useful during the pandemic and beyond.

F. Device Form, Fit and Function

The contact thermometer can be designed to take different forms, a desktop device, or a wearable device. Both devices have advantages and disadvantages. A

wearable device has potential advantages that render it suitable for the current application:

1. In one survey, pregnant women expressed willingness to use wearables during pregnancy for prenatal care [32]. In one study, the use of a wearable for self-care was demonstrated to be feasible during pregnancy [33].
2. A wearable device is portable and allows the subject to perform self-checks regardless of location. A wearable device has a physical footprint that allows for it to be worn and/or carried along by the user. The contact thermometer is intended to be used periodically, and not as a continuous monitoring device.
3. Engagement [34] in self-care and long-term behavior change [35] are potentially possible with wearable use. Women's Health, especially in disadvantaged, underserved, and vulnerable communities is a challenge and wearables can aid in alleviating the quality of life for women in such communities.
4. Wearables can be combined with smartphones to accomplish multiple objectives. Temperature screening data from the contact thermometer can be uploaded to a cloud application with the user's permission to provide data for analysis and further research. The data may be utilized in pregnancy monitoring beyond the COVID-19 mandate to assist in future efforts to alleviate morbidity and mortality during pregnancy. Smartphone applications in general may be indicative of alleviation of other conditions such as gestational diabetes mellitus [36]. Providing low-cost smartphones with multiple applications in conjunction with the wearable may accrue other pregnancy-related health benefits.
5. The team is currently progressing with research on using the wearable as an additional tool for contact tracing [37], independent of a smartphone. It may be used as a supplement to smartphone applications in contact tracing.

III. DESIGN AND CONSTRUCTION

In this section, we describe the first-generation Proof-Of-Concept (PoC) demonstration. All elements of the design, as well as future updates, are available for adoption and modification through an online repository to the scientific community through an MIT License [38].

A. Silicon Bandgap Temperature Sensor

In alignment with the design goals, the Max30205 sensor [27] is relatively inexpensive [39], has a footprint suitable for integration into a wearable form factor, usable with Printed Circuit Board (PCB) designs and is suited for fitness and clinical applications, meeting ASTM E1112 specification in the soldered version [40].

B. Microcontroller

The Xiao Microcontroller [41] based on the Arduino architecture, an open hardware architecture [42], with a list price of $4.90, has a footprint suited for prototyping low-cost designs. The ultimate solution would require a CPU with a smaller footprint, suited for wearable products. The team is currently examining alternatives for replacement.

C. Printed Circuit Board Assembly

Following the selection of key components, a minimal PCBA was designed. OSH Park [43], a low-cost, quick turnaround vendor with whom the team has experience creating other low-cost design iterations was used to manufacture the boards. Three boards were ordered and obtained for $7.55 USD each, and the components were soldered by the team. The design for the PCBA is available for adoption [38], [44]. The PCBA is defined and displayed in Fig. 1 - 4. Table 1 lists key items for the PCBA in a Bill of Materials.

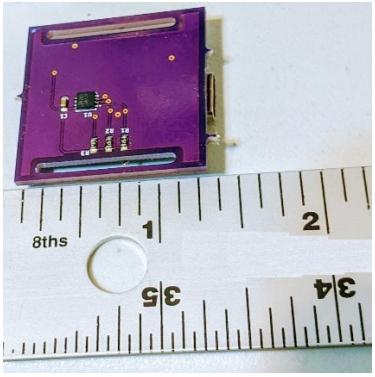


Fig. 1. Fabricated Prototype PCBA for low-cost contact thermometer (bottom view).

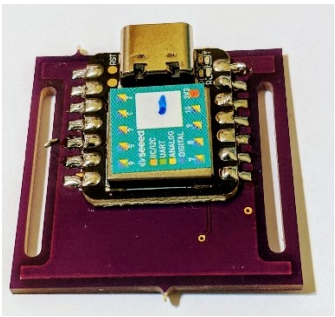


Fig. 2. Fabricated Prototype PCBA for low-cost contact thermometer (top view).

The PCBA was designed to have a small footprint, sufficient for prototyping. The current design has a square footprint, 31.75mm (1.25in) x 31.75mm (1.25in), and a thickness of 28.58mm (1.125in). The design has slots to allow straps to be inserted, and the team has recognized the slot placement, width, and strap selection will be required to be optimized in future designs. This design requires further optimization if a watch profile is

preferred when other components and measurement modalities will be added in future designs.

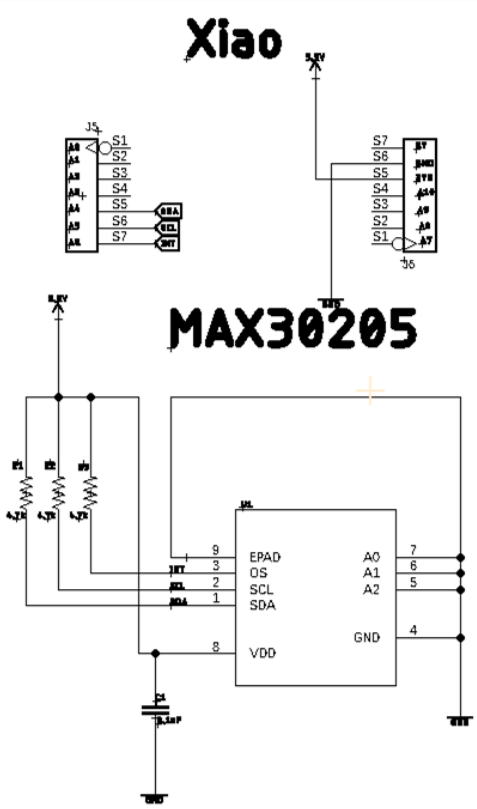


Fig. 3. PCBA schematic for low-cost contact thermometer prototype.

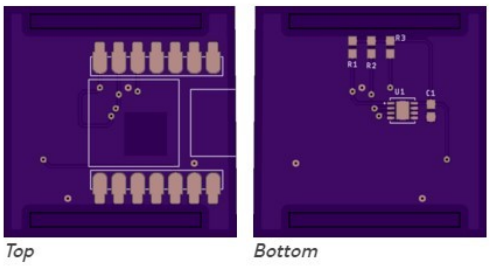


Fig. 4. PCBA Layout Diagram for low-cost contact thermometer

TABLE I. BILL OF MATERIALS - PCBA

Description		
Name	***Part Number***	***Location***
MAX30205 Temperature Sensor	MAX30205MTA+	U1
Xiao Micrcontroller	102010328	J5, J6
0.1uF capacitor	C0603C104Z3VACTU	C1
4.7k Resistor	RR0816P-472-D	R1, R2, R3

D. Other Details

The Xiao Microcontroller has a Universal Serial Bus (USB) port which was interfaced with a Personal Computer (PC) to examine the results. In future designs, this will be replaced with Bluetooth connectivity and the wearable will be paired with a smartphone application. The goal is to use coin cells, such as the CR 2032. The Max30205 sensor consumes 600uA during measurement and thus will be required to be switched off and turned on only while making measurements.

IV. DISCUSSION

A. Results

Processing [45], the standard open-source tool that can be interfaced with Arduino was used to examine the results of the prototype. The small footprint of the Max30205 produces a challenge when measuring temperature. The goal for the low-cost design is to provide screening, with the aid of surface temperatures.

Oral temperatures tend to be more accurate, however, equipment used orally requires disinfection and cleaning after every use, or the use of single-use, disposable thermometers. Armpit (axillary) and forehead (temporal) temperature measurements are expected to be 0.3°C – 0.6°C below oral temperature measurements, in value [46]. With the Max30205, axillary measurements were seen approximately 1°C below what was observed with digital thermometer measurements in a healthy subject, and approximately 2.5°C below, when measured temporally. Temporal measurements are expected to be the lowest [47]. The results observed by the team were also in line with observations made through a non-contact thermometer designed by the team, where the measurements were consistent, but below body temperature, by approximately the same range [19].

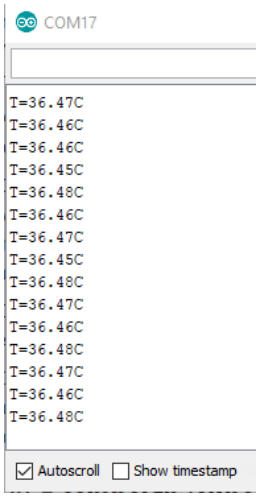


Fig. 5. Axillary Temperature Measurements using low-cost contact thermometer prototype

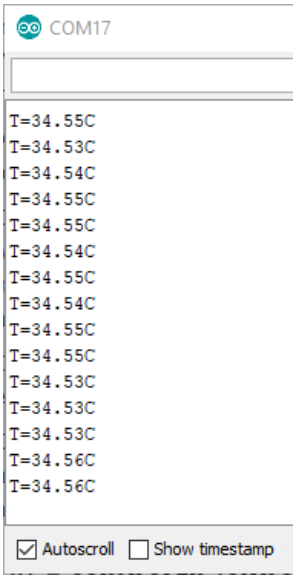


Fig. 6. Temporal Temperature Measurements using low-cost contact thermometer prototype

The measurements are not expected to be clinical, and for screening purposes in a low-cost design setting, the prototype was found to produce acceptable results. The raw data result streams are displayed in Fig. 5 and Fig. 6 for axillary and temporal measurements, respectively.

B. *Future Work*

The following work is planned for the next phases of the product design:

1) Minimize the overall footprint of the device, and design the product to have a consistent profile, such as that of a wearable and complete housing designs to prototype and test use among volunteer to iterate the design.
2) Through auditory cues such as tones, and visual cues such as light, location, and pressure as well as time to complete measurement. Temperature values or screening results must be displayed in an actionable manner.
3) Identify the appropriate chipset for processing, complete Bluetooth connections and smartphone integration. Explore contact tracing and machine learning capabilities in future application design.

V. CONCLUSION

The proof-of-concept demonstrates that a low-cost contact thermometer solution for screening is feasible. It may be applicable to aid pre-natal and post-natal care. The open design allows for multiple teams to collaborate and scale solutions as desired to combat COVID-19 and routine care during pregnancy and for other cohorts as may become necessary.

REFERENCES

[1] United Nations, “Goal 3 | Department of Economic and Social Affairs,” Accessed on July 10, 2020. [Online]. Available: https://sdgs.un.org/goals/goal3

[2] World Health Organization, “WHO | Sustainable Development Goal 3: Health,” Accessed on July 10, 2020. [Online]. Available: https://www.who.int/topics/sustainable-development-goals/targets/en/

[3] J.H. Hardie, and N.S. Landale, “Profiles of Risk: Maternal Health, Socioeconomic Status, and Child Health”, Journal of marriage and the family vol. 75,3 (2013): 651-666. doi:10.1111/jomf.12021

[4] World Health Organization, “Every Newborn: an action plan to end preventable deaths,” WHO, 2014.

[5] H. Rashid, M.Kagami, F. Ferdous, E. Ma, T. Terao, T. Hayashi, Y. Wagatsuma, "Temperature during pregnancy influences the fetal growth and birth size" Trop. Med. Health 45, 1 (2017). https://doi.org/10.1186/s41182-016-0041-6

[6] G. Tadele and A. Wasie, “Correlates of maternal mortality in developing countries: an ecological study in 82 countries,” Maternal health, neonatology and perinatology vol. 3 19. 7 Nov. 2017, doi:10.1186/s40748-017-0059-8

[7] Dynamed. “COVID-19 and Pregnant Patients,” Accessed on July 10, 2020. [Online]. Available: https://www.dynamed.com/condition/covid-19-and-pregnant-patients

[8] Health Resources & Services Administration, “Medically Underserved Areas and Populations (MUA/Ps)”, May 2020. Accessed on July 10, 2020. [Online]. Available: https://bhw.hrsa.gov/shortage-designation/muap

[9] D. Goodridge and D. Marciniuk, “Rural and remote care: Overcoming the challenges of distance,” Chronic respiratory disease vol. 13,2 (2016): 192-203. doi:10.1177/1479972316633414

[10] L.J. Green, L.H. Mackillop, D. Salvi, R. Pullon, L. Loerup, L. Tarassenko, J. Mossop, C. Edwards, S.Gerry, J.Birks, R.Gauntlett, K.Harding, L.C. Chappell and P.J.Watkinson, "Gestation-Specific Vital Sign Reference Ranges in Pregnancy," Obstetrics & Gynecology: March 2020 - Volume 135 - Issue 3 - p 653-664 doi: 10.1097/AOG.0000000000003721

[11] N. Vousden, H.L. Nathan, A.H. Shennan, "Innovations in vital signs measurement for the detection of hypertension and shock in pregnancy," Reprod Health 15, 92 (2018). https://doi.org/10.1186/s12978-018-0533-4

[12] T. Bajarin, “Why the Maker Movement Is Important to America’s Future,” Time, May 19, 2014, Accessed on July 10, 2020. [Online] Available: https://time.com/104210/maker-faire-maker-movement/

[13] CDC, “Symptoms of Coronavirus | CDC,” Accessed on July 10, 2020. [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html

[14] W. Eisenberg, “Johns Hopkins team launches temperature-tracking study and app to map and monitor potential COVID-19 cases, Accessed on July 10, 2020. [Online]. Available: https://hub.jhu.edu/2020/04/30/johns-hopkins-covid-temperature-tracking-app/

[15] San Francisco Department of Public Health, “Interim Guidance: Measuring Temperatures when Screening for COVID-19 Symptoms,” May 26, 2020.

[16] T.C. Frankel, “A fever is 100.4 in Ohio; it’s 99.5 in Delaware: States, companies write their own rules for temperature screening in a pandemic,” The Washington Post, May 15, 2020. Accessed on July 10, 2020. [Online]. Available: https://www.washingtonpost.com/business/2020/05/15/fever-screening-coronavirus/

[17] WHO, “Getting your workplace ready for COVID-19,” March 3, 2020.

[18] CDC, “Screening Clients for COVID-19 at Homeless Shelters or Encampments,” May 28, 2020. Accessed on July 10, 2020. [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/community/homeless-shelters/screening-clients-respiratory-infection-symptoms.html

[19] N.S. Yamanoor, S. Yamanoor, K. Srivastava, “Low-Cost Design of Non-Contact Thermometry for Diagnosis and Monitoring,” *2020 IEEE Global Humanitarian Technology Conference (GHTC)., to be published*.

[20] US FDA, “Non-contact Temperature Assessment Devices During the COVID-19 Pandemic,” June 19, 2020. Accessed on July 13, 2020. [Online]. Available: https://www.fda.gov/medical-devices/emergency-situations-medical-devices/non-contact-temperature-assessment-devices-during-covid-19-pandemic

[21] TEGAM, “Advantages of Contact Thermometers over Non-Contact/Infrared Thermometers,” November 14, 2015. Accessed on July 14, 2020. [Online]. Available: https://www.tegam.com/advantages-of-contact-thermometers-over-non-contactinfrared-thermometers/

[22] J. Harrison, “Choosing the Best Type of Temperature Sensor,” Maxim Integrated, August 9, 2018. Accessed on July 14, 2020. [Online]. Available: https://www.maximintegrated.com/en/design/blog/choosing-the-best-type-of-temperature-sensor.html

[23] P. Mellergård, “Intracerebral temperature in neurosurgical patients: intracerebral temperature gradients and relationships to consciousness level.” Surgical neurology vol. 43,1 (1995): 91-5. doi:10.1016/0090-3019(95)80049-m

[24] P. Narczyk, K. Siwiec and W. A. Pleskacz, "Precision human body temperature measurement based on thermistor sensor," 2016 IEEE 19th International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS), Kosice, 2016, pp. 1-5, doi: 10.1109/DDECS.2016.7482451.

[25] M. R. Neuman, "Measurement of Vital Signs: Temperature [Tutorial]," in IEEE Pulse, vol. 1, no. 2, pp. 40-49, Sept.-Oct. 2010, doi: 10.1109/MPUL.2010.937907.

[26] K. A. A.. Makinwa, “Smart temperature sensors in standard CMOS.” Procedia Engineering 5 (2010): 930-939.

[27] Maxim Integrated. “Human Body Temperature Sensor,” Accessed on July 14, 2020. [Online]. Available: https://www.maximintegrated.com/en/products/interface/sensor-interface/MAX30205.html

[28] Omega, “Non-Contact IR Human Temperature Measurement FAQs,” May 13, 2020. Accessed on July 15, 2020. [Online]. Available: https://www.omega.com/en-us/resources/body-temperature-screening-faq

[29] S. Nakamura, S.M. Ichisato, S. Horiuchi, T. Mori and M. Momoi, "Pregnant women's awareness of sensitivity to cold (hiesho) and body temperature observational study: A comparison of Japanese and Brazilian women," BMC Res Notes. 2011 Aug 5;4:278. doi: 10.1186/1756-0500-4-278. PMID: 21816112; PMCID: PMC3160981.

[30] B. A MacRae, S. Annaheim, C. M. Spengler and R.M. Rossi,"Skin Temperature Measurement Using Contact Thermometry: A Systematic Review of Setup Variables and Their Effects on Measured Values," Front. in Physiol., vol. 9, 2018. doi: 10.3389/fphys.2018.00029.

[31] Y. Freer, and A.W. Lyon, “Temperature monitoring and control in the newborn baby” Paediatrics and Child Health, 22, 127-130. doi: 10.1016/J.PAED.2011.09.002

[32] J.Runkle, M. Sugg, D. Boase, S. L. Galvin, and C. Coulson, "Use of wearable sensors for pregnancy health and environmental monitoring: Descriptive findings from the perspective of patients and providers," DIGITAL HEALTH, Jan 2019. doi: 10.1177/2055207619828220.

[33] K. Grym, H. Niela-Vilén, E. Ekholm, L. Hamari, I. Azimi, A. Rahmani, E.L. Liljeberg, and A. Axelin, "Feasibility of smart wristbands for continuous monitoring during pregnancy and one month after birth," BMC Pregnancy Childbirth 19, 34 (2019). doi: 10.1186/s12884-019-2187-9.

[34] L.A. Bove, “Increasing Patient Engagement Through the Use of Wearable Technology,” Jour. Nur. Prac., Vol. 15, Issue 8, P535-539, September 01, 2019. doi: 10.1016/j.nurpra.2019.03.018

[35] S.J. Strath andT.W. Rowley, "Wearables for Promoting Physical Activity", Clin. Chem., Vol. 64, Issue 1, 1 January 2018, Pages 53–63, doi: 10.1373/clinchem.2017.272369

[36] H. Miremberg, T. Ben-Ari, T. Betzer, H. Raphaeli, R. Gasnier, G. Barda, J. Bar and E. Weiner, "The impact of a daily smartphone-based feedback system among women with gestational diabetes on compliance, glycemic control, satisfaction, and pregnancy outcome: a randomized controlled trial," American journal of obstetrics and gynecology, 218(4), 453.e1–453.e7. doi: 10.1016/j.ajog.2018.01.044

[37] J. Hellewell, S. Abbott, A. Gimma, N.I. Bosse, C.I. Jarvis, T.W. Russell, J.D. Munday, A.J. Kucharski, W.J. Edmunds, S. Funk and R.M. Eggo, "Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts," Lancet Glob Health. 2020;8(4):e488 – e496.

[38] N.S. Yamanoor and S. Yamanoor, “GitHub – sai-ydev/max30205.” Accessed on August 05, 2020. [Online]. Available: https://github.com/sai-ydev/max30205

[39] Digikey, “MAX30205 – Maxim Integrated – Temperature – Analog and Digital Output | Online Catalog | DigiKey Electronics.” Accessed on July 25, 2020. [Online]. Available: https://www.digikey.com/catalog/en/partgroup/max30205/65060

[40] *Standard Specification for Electronic Thermometer for Intermittent Determination of Patient Temperature.* ASTM E1112-00. 2018.

[41] Seeed Studio, “Seeeduino XIAO - Arduino Microcontroller - SAMD21 Cortex M0+.” Accessed on July 25, 2020. [Online]. Available: https://www.seeedstudio.com/Seeeduino-XIAO-Arduino-Microcontroller-SAMD21-Cortex-M0+-p-4426.html

[42] A. Goel, “Arduino (ATmega328/p) Architecture Explained,” Eckovation, May 17, 2018. Accessed on July 28, 2020. [Online]. Available: https://engineering.eckovation.com/arduino-architecture-explained/#:~:text=Architecture%3A%20means%20the%20complex%20or,reduced%20instruction%20set%20computer)%20architecture.

[43] OSH Park, “Our Services.” Accessed on July 15, 2020. [Online]. Available: https://oshpark.com/#services

[44] N.S. Yamanoor and S. Yamanor, “max30205.” Accessed on July 20, 2020. [Online]. Available: https://oshpark.com/shared_projects/WVVYrYFY

[45] Arduino, “Arduino Playground – Processing.” Accessed on July 25, 2020. [Online]. Available: https://playground.arduino.cc/Interfacing/Processing/

[46] C.S. Mott Children’s Hospital “Fever Temperatures: Accuracy and Comparison.” Accessed on July 28, 2020. [Online]. Available: https://www.mottchildren.org/health-library/tw9223#:~:text=An%20armpit%20(axillary)%20temperature%20is,lower%20than%20an%20oral%20temperature.

[47] S. Asadian, A. Khatony, G. Moradi, A. Abdi and M. Rezaei, "Accuracy and precision of four common peripheral temperature measurement methods in intensive care patients," Med Devices (Auckl). 2016;9:301-308. Published 2016 Sep 1. doi:10.2147/MDER.S109904

EPD-Net: A GAN-based Architecture for Face De-identification from Images

Alakh Aggarwal
CSE Department
IIT (BHU), Varanasi
Varanasi, India
alakh.aggarwal.cse15@iitbhu.ac.in

Rishika Rathore
Chemical Engineering Department
IIT (BHU), Varanasi
Varanasi, India
rishika.rathore.che17@iitbhu.ac.in

Pratik Chattopadhyay
CSE Department
IIT (BHU), Varanasi
Varanasi, India
pratik.cse@iitbhu.ac.in

Lipo Wang
School of EEE
Nanyang Technological University
Singapore
ELPWang@ntu.edu.sg

***Abstract*—Nowadays huge amount of crowd data captured by surveillance cameras gets shared publicly in the form of images or videos through television or the internet. Although many of these videos are meant to provide public security, they also lead to a widespread concern towards privacy protection, since a lot of personal information about subjects gets revealed through this video/image data. Hence, "de-identifying" people (i.e., obscuring identity information) captured by the surveillance cameras is of utmost importance for providing privacy along with security. Traditional identity obfuscation techniques such as blurring, warping, and filtering lead to the loss of vital non-biometric information. More recent k-same-based as well as generative model-based de-identification techniques eliminate the above problem to a certain extent. Still, the visual quality of the generated images produced by these methods is not realistic. Also, these approaches are unable to maintain structural integrity and cannot preserve the required non-biometric information at a high resolution. As an improvement, in this paper, we propose a new network termed as EPD-Net and train it with suitable loss functions to maximize the emotion similarity and minimize the identity similarity. Experimental results verify the effectiveness of our approach and its superiority over other popular face de-identification techniques.**

***Index Terms*—De-identification, Identity Obfuscation, Emotion Preservation, Generative Adversarial Networks, Privacy Preservation**

I. Introduction

In the present days, a large volume of data gets shared over the internet across different parts of the world. In this process of transfer and share of visual data, leak of privacy becomes the primary concern, since these data reveal the biometric traits of the subjects without their consent. Most real-life applications working on crowd data focus on providing public security, where knowledge of the identity of individuals has lower significance, and capturing other non-biometric information such as race, gender, facial expression/emotion have a higher significance. For example, consider a public hospital scenario where the number of patients is large, in comparison to the number of hospital staff, and monitoring the condition of every patient constantly becomes very difficult. The facial expression of patients captured by the surveillance cameras installed in the hospital corridors and/or patient cabins provide vital clues about the physical and mental conditions of the patients. By observing this data the hospital authorities can take appropriate measures and/or provide necessary treatments, if required. However, these videos often reveal the identity of the patients captured in the videos. From the above discussion, it is clear that there is a need for de-identifying patients' identity, while simultaneously preserving their emotions. To the best of our knowledge, there exists no work in the literature that solves this problem effectively except the work in [1].

Fig. 1: Sample images from RafD (Radboud Faces Dataset) consisting of 67 subjects belonging to different classes, namely, Caucasian male, Caucasian female, Moroccan male, Kids male, Kids female, etc.

978-1-7281-9615-2/20/$31.00 ©2020 IEEE

Although the objective of this work is similar to that of ours, the images generated by this method lack in quality, and the identifiable characteristics of a face are not eliminated well. The de-identified faces can be easily identified visually and this makes it unsuitable for application in real-life.

In this paper, we aim to improve the existing solutions to the problem of face de-identification, while simultaneously preserving emotion (or, facial expression) information. GANs [2] have been extensively used in recent years for several image generation or translation tasks such as image de-fencing [3], colorization [4], etc. The present work is an application of image to image translation, where a facial image is provided as input, and a de-identified version of the same face is obtained as the output. Motivated by the recent success of conditional GANs (in short, cGANs) in image translation tasks [5], we also use this category of the network in the present work. Specifically, we propose a new network architecture termed as EPD-Net (Emotion Preserving De-identification Network), that consists of a pix2pix GAN [5] along with two additional auxiliary deep networks, namely the identity and emotion verificators. When a face image is input to the EPD-Net, the network does certain computations in its layers to generate a realistic de-identified face image that preserves the emotional characteristics of the input face and eliminates identity characteristics to the extent possible.

The rest of the paper is organized as follows. A thorough survey of related literature is discussed in Section II. Section III introduces the proposed EPD-Net network along the loss functions used in training the model. Extensive experimental evaluation and comparison with some competing approaches are presented in Section IV. Section V finally concludes the paper and points out possible future scopes for research.

II. Related Work

Privacy protection in images/videos is very important since most online images and surveillance videos inadvertently disclose the identities of subjects captured. In the past, research work on de-identification has been mostly focused on developing algorithms for facial identity obfuscation such as [6]–[12], and only a few gait/silhouette de-identification techniques have been proposed, such as [13], [14]. The gait/silhouette de-identification algorithms commonly use only blurring operator to de-identify a subject silhouette, which cannot always guarantee anonymity. Since the theme of the paper is developing algorithms on face de-identification by preserving emotional characteristics, here we focus on describing the research trend on face de-identification, and skip discussions related to gait/silhouette de-identification.

Previous techniques used for de-identification were simplistic in nature. They were mostly aimed at concealing detailed texture information present on the face by the means of blurring [6], warping [7], [8], pixelation [15], and other image processing-based techniques. This category of these approaches is termed as non-formal, since these do not provide any formal proof of anonymity. As an improvement, k-same family of approaches was developed later, which has been used for de-identification tasks in various studies such as [9]–[11]. These approaches, on the surface, work by finding cluster centers of k-clusters from the data set and replacing the cluster instances with the cluster centers. Although the drawback of uncertain anonymity can be successfully handled by the means of the k-same family of approaches, the resulting de-identified images look highly non-realistic, and bear ghostly appearances, due to the averaging operation used in computing the cluster centers. With the introduction of Generative Neural Networks (GNNs), learning based techniques were developed to generate better quality de-identified faces, for example, [1], [12].

In [12], Meden et. al. proposed a generative neural network (GNN) based approach in which proxy clusters are generated using the k-same algorithm and the de-identified faces are generated by the GNN. The complete de-identification network has been termed as the k-same-net, and although it improves over the above-mentioned non-formal techniques, the generated images still look non-realistic. Moreover, each of the k-same family of de-identification approaches, including k-same-net, suffer from the following two shortcomings: (i) The k-same approach assume that each subject is only represented once in the data set, but this may not be true in practice. The presence of multiple images from the same subject, or images sharing similar biometric characteristics can lead to lower levels of privacy protection, (ii) k-same operates on a closed set of facial images and produces a corresponding de-identified set, which is not applicable in situations that involve the processing of individual images, or sequence of images outside the domain of data set.

More recent approaches employ the generalization power of Generative Adversarial Networks (GANs) [2] to de-identify faces. The basic architecture of a GAN consists of a generator network that generates an output based on the input data's probability distribution function and a discriminator that distinguishes the generated output from the input data. Typically, a GAN is trained in multiple iterations using Minimax algorithm, and this imparts an ability to the generator network to accurately fit the input probability distribution function. In recent years, researchers have aided the generation and discrimination process of GANs with slightly different but similar concepts. For example, the work in [1] employs a GAN along with a verificator network to obtain a similarity measure between the features of the generated image and the desired features, as well as a regularizer network to compute the structural similarity between the two images. The objective function for training the GAN accounts for the combined loss given by the verificator and the discriminator networks. However, the objective function in [1] does not take into account emotion preservation. Hence, the de-identified images generated by this approach cannot guarantee emotion preservation effectively. Our observation is that the images generated by [1] also do not look significantly realistic (results are shown in Section IV).

In this paper, we propose an improvement to the work in [1], by adding an additional verificator network to preserve

the emotion information at a higher resolution compared to [1]. Also, we propose to use a deep classification network as an identity verificator network instead of the Siamese network as in [1]. This is since the function of the identity verificator is to predict if the generated face corresponds to a particular identity. This can be effectively done if the network is trained to learn a mapping from an input face to its corresponding identity. The Siamese network, on the other hand, does not take into account the ground truth label information, and predicts the similarity of the generated face with the original face based on the similarity in their appearances. The use of the classification network as an identity verificator also helps in obfuscating identity information better than that [1].

The main contributions of the work are as follows:

- Developing an improved network termed as EPD-Net for face de-identification that preserves emotion characteristics of a face effectively while simultaneously obfuscating the facial identity information.
- Improving the identity verificator network used in a previous study [1] by employing a deep convolution classification network instead of Siamese network.
- Extensive experimental evaluation and comparison with competing approaches.

III. Proposed Approach

The face de-identification process can be viewed as a mapping from an input face image x to its corresponding de-identified image $\hat{x}$ using a mapping function f such that $f(x) = \hat{x}$. In an ideal situation, the mapping function f must be non-invertible, i.e., it should not be possible to recover the original face from the de-identified face. But as discussed in Section I, obfuscation of identity features is not the only goal of this work. Along with identity obfuscation, emotion loss should be minimized to the extent possible. Generative Adversarial Networks (GANs) are a popular choice for researchers to generate images following a specific distribution pattern. Since the objective of the present work is generating face images by satisfying a given set of conditions, it appears that a suitable GAN-based architecture can be employed to learn the mapping function f. The basic version of GAN consists of two networks: (a) a generator network which tries to learn the input data distribution, and (b) a discriminator network which predicts whether an input image belongs to the input distribution (i.e., real data) or the distribution learned by the generator network (i.e., fake data). However, using this basic structure of GAN, it is not possible to achieve the dual objective of identity obfuscation and emotion preservation. If a set of auxiliary features needs to be employed while simultaneously learning the input data distribution, then conditional GANs (in short cGANs) [16] seem to be the best choice. In a cGAN, separate loss functions are designed corresponding to each auxiliary information using either standard metrics or neural networks, and the combined loss from all the auxiliary networks/metrics are jointly minimized along with the discriminator loss function.

To construct the EPD-Net, we consider two separate deep networks along with a pix-2-pix GAN and design loss functions to satisfy the following conditions: (a) the generated image must be free from identifiable features present in the input face, and (b) the generated image must preserve the emotion of the input face. We term these auxiliary networks as *identity verificator* and *emotion verificator* networks, and these produce scores depicting the identity similarity and emotion dissimilarity between the input and the generated de-identified faces, respectively. The architecture of the proposed network can be schematically represented using Figure 2. With reference to the figure, the network has the following

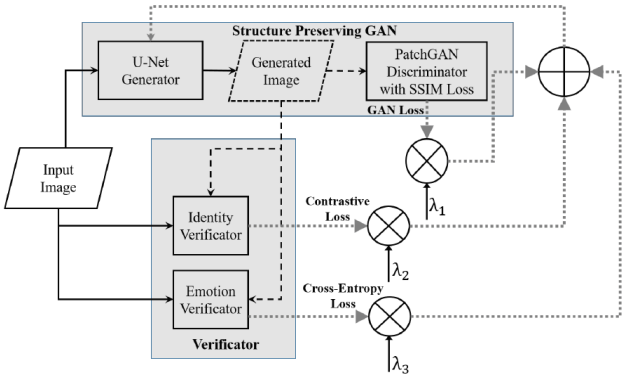


Fig. 2: An Overview of the proposed network architecture

modules: (a) a cGAN that is based on pix-2-pix network as described in [5], and (b) identity and emotion verificator networks, as described later. The module (a) consists of a generator and a discriminator. The generator generates an image and the discriminator outputs a score (shown as *GAN Loss*) depicting whether the generated image belongs to the probability distribution of the gallery set or not. The verificator networks compute two loss metrics, shown as *Contrastive Loss* and *Cross-Entropy Loss*, that respectively indicate the extent of identity similarity and emotion dissimilarity between the input and the generated images. A weighted combination of these loss values at a particular epoch is fed back to the generator, which then generates an improved face image in the subsequent epoch. The combination tallies more accurately with the different conditions or constraints. In the figure, λ_1, λ_2, and λ_3 represent the weighting factors corresponding to *GAN Loss*, *Contrastive Loss*, and *Cross-Entropy Loss*, respectively.

A. Structure Preserving GAN

The GAN architecture and the training algorithm used for face de-identification are described in detail next. Unlike the previous work on face de-identification, we focus on generating more realistic face images with distinct facial features (i.e., eye, nose, mouth, etc.) and clear face boundaries. For this, a structural similarity loss term is also optimized along with the standard discriminator loss functions in the objective function. The pix-2-pix GAN consists of an U-Net generator

[17] and a variant of the PatchGAN discriminator, which is a fully convolutional network formed by the *Conv-BatchNorm-ReLU* blocks as described in [5]. The loss function for training the model is computed at the final layer of this discriminator. The generator is fed with a random noise vector (say z), derived from noise distribution given by P_z. Let us also assume that the input data distribution is given by P_{data}, and the distribution learned by the generator at a particular epoch by P_{gen}. For each input image x in the training set, the generator performs a mapping of the form $G : (x, z) \mapsto \hat{x}$, where $x \in P_{data}$ and $\hat{x} \in P_{gen}$. The GAN is trained in multiple epochs and the training process stops when the two distributions P_{gen} and P_{data} become similar.

As the generator continues to learn the input distribution in a better way at each epoch, the discriminator network also keeps on improving its prediction power so that it can correctly classify a given image into the appropriate class: i.e., real or fake. As already explained before, this is a binary classification task where the two classes correspond to samples from the distributions P_{data} and P_{gen}. Let us assume that the ground truth label corresponding to P_{data} is [0,1] while for P_{gen} it is [1,0]. The mutual conflict between the two networks helps in gradual improvement of the distribution function learned by the generator, as well as the prediction capability of the discriminator. On completion of the training phase, the generator learns the distribution P_{data} accurately and it attains the capability of generating images that look similar to the original data.

Suppose, at a particular epoch, the discriminator is trained with N training patterns, and let the i^{th} pattern be denoted by y_i, where y_i is the input image. The discriminator distinguishes the probability distribution in which y_i belongs, such that $y_i \in P_{data}$ or $y_i \in P_{gen}$. The mini-max loss function to be optimized for the generator-discriminator network can be represented by (1):

$$\begin{aligned} &L_{PatchGAN}(\{y_1, y_2, ...y_N\}, D(x, y)) = \\ &\mathbf{E}_{x,\hat{x}\in P_{data}(x,\hat{x})}[log(D(x, \hat{x}))]+ \\ &\mathbf{E}_{x\in P_{data}(x,\hat{x}),z\in P_z(z)}[log(1 - D(x, G(x, z)))], \end{aligned} \tag{1}$$

where $\mathbf{E}$ denotes the expectation operator. We observe that training the cGAN with the loss function of (1) alone fails to preserve the emotion and facial structure effectively. The rendered image often looks unrealistic and ghostly. To preserve the above non-biometric features at a high resolution, and generate an aesthetically pleasing de-identified face, we add a structural similarity loss term (denoted by $L_{ssim}(x, \hat{x})$ [18]) between the input face (x) and the generated face ($\hat{x}$) along with the loss function of (1). This loss term is computed by taking into consideration the similarities in the contrast, luminance and structure between x and $\hat{x}$, and is a combined measure of these three. If the luminance, contrast and structure similarities are respectively denoted by $l(x, \hat{x})$, $c(x, \hat{x})$, and $l(s, \hat{x})$, then the structural similarity index (SSIM) [19] between two input images x and $\hat{x}$ is mathematically defined as

$$SSIM(x, \hat{x}) = l(x, \hat{x})^{\alpha}.c(x, \hat{x})^{\beta}.s(x, \hat{x})^{\gamma}, \tag{2}$$

where,

$$l(x, \hat{x}) = \frac{2\mu_x\mu_{\hat{x}} + c_1}{\mu_x^2 + \mu_{\hat{x}}^2 + c_1}, \tag{3}$$

$$c(x, \hat{x}) = \frac{2\sigma_x\sigma_{\hat{x}} + c_2}{\sigma_x^2 + \sigma_{\hat{x}}^2 + c_2}, \tag{4}$$

$$\text{and } s(x, \hat{x}) = \frac{\sigma_{x\hat{x}} + c_3}{\sigma_x\sigma_{\hat{x}} + c_3}. \tag{5}$$

In the above expressions, μ_x and $\mu_{\hat{x}}$ denote the average intensities of the inputs x, $\hat{x}$, σ_x^2 and $\sigma_{\hat{x}}^2$ refer to the variances in the intensities of x and $\hat{x}$, and $\sigma_x\sigma_{\hat{x}}$ denotes the covariance of x and $\hat{x}$. Constants c_1, c_2 and c_3 are defined as follows: c_1=$(k_1L)^2$, c_2=$(k_2L)^2$, c_3=$c_2/2$, k_1=0.01, k_2 = 0.03 where L is the dynamic range of the pixel values. For example, the value of L is 255 for 8 bit gray-level images. The structural similarity loss metric $L_{ssim}(x, \hat{x})$ is computed from (2) according to the following expression:

$$L_{ssim}(x, \hat{x}) = \frac{1}{2}(1 - SSIM(x, \hat{x})). \tag{6}$$

If L_{cGAN} denotes the loss function to be minimized for the generator and the discriminator networks without enforcing any condition, then L_{cGAN} can be mathematically represented as follows:

$$L_{cGAN}(x, \hat{x}, D(x, \hat{x})) = L_{PatchGAN}(x, \hat{x}) + \Phi L_{ssim}(x, \hat{x}). \tag{7}$$

where, Φ is a positive constant in the range [0,1]. Since there is a trade-off between effective de-identification and structural similarity preservation, the value of Φ must be chosen carefully to maintain a proper balance between the two. We observe that a value of Φ equal to 0.25 can help in achieving the desired balance.

B. Verificator Networks

As shown in Figure 2, we employ two different verificator networks to minimize identity similarity and maximize emotion similarity. These networks are discussed next.

1) Identity Verificator: VGG-16 Convolutional Neural Network [20] is employed as an identity verificator which takes as input the original image x and the generated image $\hat{x}$ and computes the similarity in the identity features between these images. Let us consider that there are M faces in the gallery set and the VGG-16 identity verificator network has M output nodes corresponding to each identity. If I_x denotes the ground-truth identity label for person x, (where x = 1, 2, 3, ..., M), then I_x is a feature vector of dimension M, where

$$I_x(i) = \begin{cases} 1 & \text{if } i = x \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The network is initially trained on the gallery set to learn the mapping function: $V : x \mapsto I_x$. Once this verificator is trained accurately, its weights are freezed, and next this trained

network is used in conjunction with the cGAN (discussed in Section III-A) to train the generator to produce de-identified faces.

Let us consider that at a particular iteration, the generator outputs image $\hat{x}$ for a certain input x. With reference to Figure 2, $\hat{x}$ is given as input to the identity verificator network along with the ground truth identity label I_x. As a response to input $\hat{x}$, the identity verificator outputs $V_{\hat{x}}$, (where, $V_{\hat{x}} = V(\hat{x})$, i.e., it is the vector formed by concatenating the values at the output nodes on presentation of input $\hat{x}$). In any de-identification task, the objective is to maximize the distance between I_x and predicted identity features $V_{\hat{x}}$. To do this, a loss function (denoted as $L_{contrastive}$) is computed at the output layer of the identity verificator. $L_{contrastive}$ is mathematically defined as follows:

$$L_{contrastive}(V_{\hat{x}}, I_x) = max(0, \alpha - ||I_x - V_{\hat{x}}||_2)^2, \quad (9)$$

where α is a regulatory parameter. The expression (9) ensures that the value of $L_{contrastive}$ lies between 0 and α.

2) Emotion Verificator: Similar to the identity verificator network (described in Section III-B1), we employ Xception Convolutional Neural Network [21] model to find the similarity of the emotion features between the input and the generated face images. Let S denote the mapping function learned by this network for mapping an image x to its corresponding emotion class represented by E_x. When the generated image $\hat{x}$ corresponding to input x is fed to the same Xception network, it generates an output vector $S(\hat{x})$. Now, if $L_{crossentropy}(S(\hat{x}), E_x)$ denotes the loss function associated with the emotion verificator, then a measure of dissimilarity between the two emotions is obtained by applying a cross-entropy loss function as shown in (10):

$$L_{crossentropy}(S(\hat{x}), E_x) = -(E_x log(S(\hat{x})) + (1 - E_x) log(1 - S(\hat{x})). \quad (10)$$

C. De-identification with Emotion and Facial Structure Preservation

Combining all the losses with weighted-sum, we get our final objective function

$$\begin{aligned} &L_{face}(x, G(x,z), D(x, G(x,z))) = \\ &\lambda_1 L_{cGAN}(x, G(x,z), D(x, G(x,z))) + \\ &\lambda_2 L_{contrastive}(V(G(x,z)), I_x) + \\ &\lambda_3 L_{crossentropy}(S(G(x,z)), E_x), \end{aligned} \quad (11)$$

where λ_1, λ_2, λ_3 are pre-defined constants ($0 \leq \lambda_1, \lambda_2, \lambda_1 \leq 1$), and the optimal generator output G^* is found using the minimax function as follows:

$$G^* = arg \min_G \max_D L_{face}(x, G, D). \quad (12)$$

IV. Experiments and Results

A. Data Set and System Description

Our algorithm has been implemented on a system with 64 GB RAM, one i9-18 core processor, and three GPUs, out of which one is Titan Xp with 12 GB RAM, 12 GB frame-buffer memory and 256 MB BAR1 memory, and the other two are GeForce GTX 1080 Ti with 11 GB RAM, 11 GB frame-buffer memory and 256 MB of BAR1 memory.

We evaluate our algorithm on two popularly used data sets, namely the Radboud Faces Database (RaFD) [22] and Kaggle emotion detection data set. The RaFD data consists of 8040 facial images corresponding to 67 subjects belonging to five categories: Caucasian male, Caucasian female, Caucasian kid male, Caucasian kid female, and Moroccan Dutch male. For each subject, the RaFD data preserves different emotions including anger, disgust, fear, happiness, sadness, neutral, etc. Each emotion is accompanied by three gaze directions and all the snapshots are taken from five different angles. The emotion verificator model described in Section III-B is trained with the RaFD data set to detect the following emotions: anger, disgust, fear, happiness, sadness, surprise, and neutral. The learning rates used for training the Identity and Emotion verificator networks are set to 1e-6 and 1e-4, respectively.

To evaluate the effectiveness of our approach, we make a thorough experimental analysis and also perform a comparative evaluation with two popular state-of-the-art face de-identification methods (namely, PPGAN [1] and K-SameNet [12]). In Figure 3, we show some results obtained by applying the proposed de-identification algorithm on sample face images from the RaFD data. The first row in the figure represents the original face images, while the second row shows the corresponding de-identified faces.

It can be visually observed from the figure that in addition to obfuscation identity by a considerable extent, the EPD-Net also successfully preserves the emotion characteristics on the faces.

Before presenting further quantitative results, we define two metrics that have been used to evaluate the effectiveness of de-identification algorithms, namely, the de-identification rate and the switching de-identification rate. De-identification rate refers to the percentage of subjects for which identity obfuscation in the generated image is successful, whereas switching de-identification rate refers to the percentage of subjects for which the generated face's identity fails to match the identity of any subject in the gallery set. For effective de-identification, both de-identification and switching de-identification rates must be high. These metrics can be computed from the VGG-16 output score by considering a threshold parameter (say, μ), such that a verification score greater than μ implies de-identification (or, switching de-identification) is successful. We observe that a value of μ=0.35 results in a self de-identification rate of 69.90%, and a switching de-identification rate of 86.08%, which can be regarded as significantly good de-identification performance.

Next, we perform a comparative performance analysis of our approach with PPGAN [1] and K-SameNet [12] in terms of face recognition misclassification rate. To study the effect of using the SSIM loss term during training the de-identification model, in the same experiment we have also evaluated the performance of our approach without considering the SSIM

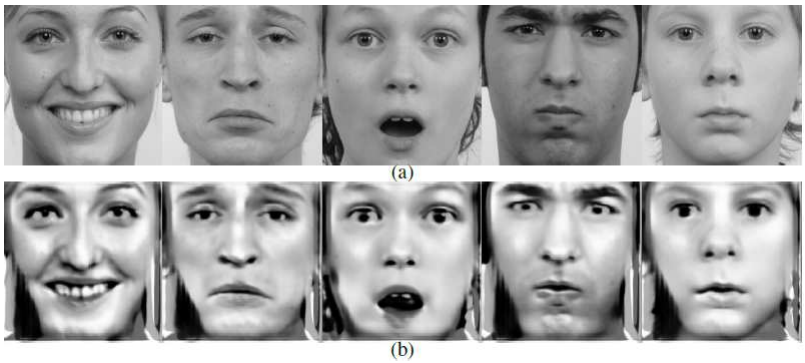


Fig. 3: (a) Sample Face Images from RaFD data, (b) corresponding de-identification results obtained by applying our approach

loss term. Results are shown in Figure 4 for ranks 1 to 5, by means of Cumulative Match Characteristic (CMC) curves. The prediction of the VGG-16 network is employed to plot these curves, i.e., if the class predicted by the VGG-16 network is different from the actual class, then misclassification occurs. It is imperative that the algorithm with higher misclassification rate should be regarded as a better de-identification algorithm. It is observed from the figure that the proposed method surpasses all the other approaches used in the comparative study in terms of misclassification rate for the different rank values. Although Rank 1 misclassification rate for each approach is similar, our approach outperforms the other methods as well as the approach without employing the SSIM loss function for each of the higher rank values. This proves that the identity obfuscation potential of the proposed approach is superior to the competing approaches.

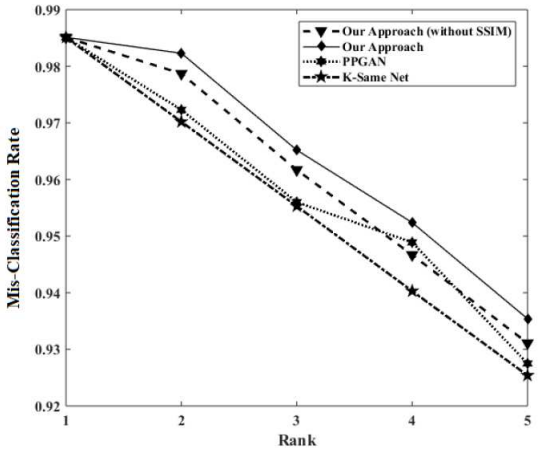


Fig. 4: Rank based mis-classification rate for our approach (with and without SSIM), PPGAN and K-Same Net

Next, we compare the performances of the different methods used in the previous experiment in terms of emotion preservation capability. Results are shown in Table I by means of averaged mean-squared error metric. For each test face image, we find the emotion features by passing the image through the Xception network, and next compute the squared difference between the two feature vectors. Finally, the mean of the square-differences from all the faces is presented in Table I. As expected, it is observed that due to the inclusion of the emotion verificator network in the complete architecture (refer to Figure 2), the generated image is able to preserve the emotion of the input image at a higher resolution compared to other state-of-the-art approaches, thereby resulting in a lower mean-squared error-rate. It can also be seen from the table that the mean-squared error of our approach increases by a factor of 0.0142 when SSIM loss is not included while training the GAN (refer to Equation 7). Thus, it can be concluded that apart from generating faces with proper facial features, the SSIM loss term also significantly helps in preserving the emotion present in the input face.

Method	Mean-Squared Error
Our Approach	0.0099
Our Approach (without SSIM)	0.0241
PPGAN	0.0257
K-same	0.0238

TABLE I: Emotion mean-squared error

In the next experiment, we study the extent to which structural similarity is preserved in case of the different approaches. Results are shown in Figure 5, by means of bar diagrams in which each bar represents the average SSIM index obtained after comparing the generated image with the original image for all the faces. It can be seen from the figure that, PPGAN [1] has marginally better performance (by a factor 0.001) compared to our approach. This is since while minimizing the identity similarity, the proposed network also obscures the color and contrast information present in the input face to a certain extent, which accounts for a slightly low SSIM score. Despite the marginally better performance of [1] over our approach in terms of SSIM score, its emotion preservation capability, as well as misclassification rates are inferior compared to our approach as already observed from Figure 4 and Table I. Thus, in terms of effective de-identification with emotion preservation, our method performs the best among

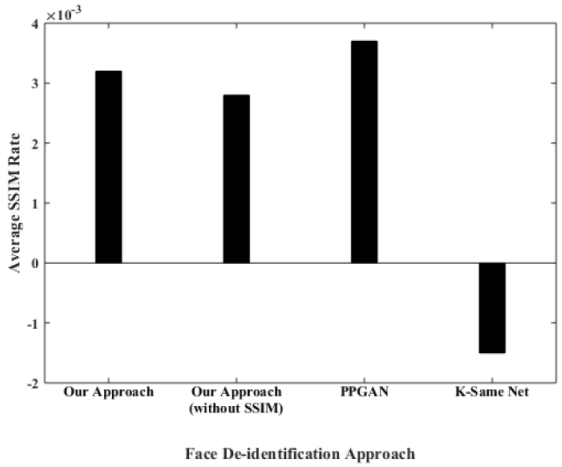


Fig. 5: Average SSIM rate given by the different face de-identification approaches

other existing learning-based approaches.

From the thorough experimental analysis, it can be concluded that the proposed approach performs de-identification effectively and has superior performance compared to other state-of-the-art face de-identification techniques. The present work can be easily extended to perform de-identification with the preservation of other non-biometric features like gender, skin-tone, race, etc., by adding separate verificator networks.

V. Conclusions and Future Work

In this paper, we propose a new architecture for face de-identification using conditional Generative Adversarial Network. The network successfully preserves emotion, but obscures the identifiable characteristics in a given face image. Extensive experimental evaluation show that our approach outperforms other state-of-the-art face de-identification algorithms in terms of identity obfuscation and emotion preservation capabilities. Although we have worked with emotion preservation only, our approach can be conveniently extended to retain any required non-biometric information along with simultaneous removal of biometric features. In future, the work can be extended to perform face de-identification in videos. The proposed de-identification technique will have significant use in commercial applications, e.g., study of customer emotional feedback for products, whereby, maintaining privacy, behavioral-based insider threat detection, and several others.

Acknowledgements

The authors also express their sincere gratitude to NVIDIA for supporting their research with a Titan Xp GPU.

References

[1] Y. Wu, F. Yang, and H. Ling, "Privacy-protective-gan for face de-identification," *arXiv preprint arXiv:1806.08906*, 2018.

[2] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, pp. 2672–2680, 2014.

[3] K. Nazeri, E. Ng, T. Joseph, F. Qureshi, and M. Ebrahimi, "Edgeconnect: Generative image inpainting with adversarial edge learning," *arXiv preprint arXiv:1901.00212*, 2019.

[4] Y. Cao, Z. Zhou, W. Zhang, and Y. Yu, "Unsupervised diverse colorization via generative adversarial networks," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 151–166, Springer, 2017.

[5] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *arXiv preprint*, 2017.

[6] A. Frome, G. Cheung, A. Abdulkader, M. Zennaro, B. Wu, A. Bissacco, H. Adam, H. Neven, and L. Vincent, "Large-scale privacy protection in google street view.," in *ICCV*, pp. 2373–2380, 2009.

[7] H. Sohn, W. De Neve, and Y. M. Ro, "Privacy protection in video surveillance systems: Analysis of subband-adaptive scrambling in jpeg xr," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 2, pp. 170–177, 2011.

[8] L. E. Schulz and E. B. Bonawitz, "Serious fun: preschoolers engage in more exploratory play when evidence is confounded.," *Developmental psychology*, vol. 43, no. 4, p. 1045, 2007.

[9] R. Gross, L. Sweeney, F. De la Torre, and S. Baker, "Model-based face de-identification," in *null*, p. 161, IEEE, 2006.

[10] R. Gross, E. Airoldi, B. Malin, and L. Sweeney, "Integrating utility into face de-identification," in *International Workshop on Privacy Enhancing Technologies*, pp. 227–242, Springer, 2005.

[11] R. Gross, L. Sweeney, F. De La Torre, and S. Baker, "Semi-supervised learning of multi-factor models for face de-identification," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pp. 1–8, IEEE, 2008.

[12] B. Meden, Ž. Emeršič, V. Štruc, and P. Peer, "k-same-net: k-anonymity with generative deep neural networks for face deidentification," *Entropy*, vol. 20, no. 1, p. 60, 2018.

[13] P. Korshunov and T. Ebrahimi, "Using warping for privacy protection in video surveillance," in *Digital Signal Processing (DSP), 2013 18th International Conference on*, pp. 1–6, IEEE, 2013.

[14] P. Agrawal and P. Narayanan, "Person de-identification in videos," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 3, pp. 299–310, 2011.

[15] O. M. Parkhi, A. Vedaldi, A. Zisserman, *et al.*, "Deep face recognition.," in *bmvc*, vol. 1, p. 6, 2015.

[16] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *CoRR*, vol. abs/1411.1784, 2014.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *CoRR*, vol. abs/1505.04597, 2015.

[18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 13, no. 4, pp. 600–612, 2004.

[19] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, *et al.*, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[21] O. Arriaga, M. Valdenegro-Toro, and P. Plöger, "Real-time convolutional neural networks for emotion and gender classification," *arXiv preprint arXiv:1710.07557*, 2017.

[22] O. Langner, R. Dotsch, G. Bijlstra, D. H. Wigboldus, S. T. Hawk, and A. Van Knippenberg, "Presentation and validation of the radboud faces database," *Cognition and emotion*, vol. 24, no. 8, pp. 1377–1388, 2010.

IEMTRONICS 2020

International IOT, Electronics and Mechatronics Conference