

Lecture Notes in Electrical Engineering 756

Saad Mekhilef
Margarita Favorskaya
R. K. Pandey
Rabindra Nath Shaw *Editors*

Innovations in Electrical and Electronic Engineering

Proceedings of ICEEE 2021

 Springer

Lecture Notes in Electrical Engineering

Volume 756

Series Editors

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Naples, Italy

Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico

Bijaya Ketan Panigrahi, Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India
Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, Munich, Germany

Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China

Shanben Chen, Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

Rüdiger Dillmann, Humanoids and Intelligent Systems Laboratory, Karlsruhe Institute for Technology, Karlsruhe, Germany

Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China

Gianluigi Ferrari, Università di Parma, Parma, Italy

Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain

Sandra Hirche, Department of Electrical Engineering and Information Science, Technische Universität München, Munich, Germany

Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA

Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Alaa Khamis, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt

Torsten Kroeger, Stanford University, Stanford, CA, USA

Yong Li, Hunan University, Changsha, Hunan, China

Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA

Ferran Martín, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore

Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany

Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA

Sebastian Möller, Quality and Usability Laboratory, TU Berlin, Berlin, Germany

Subhas Mukhopadhyay, School of Engineering & Advanced Technology, Massey University, Palmerston North, Manawatu-Wanganui, New Zealand

Cun-Zheng Ning, Electrical Engineering, Arizona State University, Tempe, AZ, USA

Toyoaki Nishida, Graduate School of Informatics, Kyoto University, Kyoto, Japan

Federica Pascucci, Dipartimento di Ingegneria, Università degli Studi "Roma Tre", Rome, Italy

Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Gan Woon Seng, School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore, Singapore

Joachim Speidel, Institute of Telecommunications, Universität Stuttgart, Stuttgart, Germany

Germano Veiga, Campus da FEUP, INESC Porto, Porto, Portugal

Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China

Junjie James Zhang, Charlotte, NC, USA

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

More information about this series at <http://www.springer.com/series/7818>

Saad Mekhilef · Margarita Favorskaya ·
R. K. Pandey · Rabindra Nath Shaw
Editors

Innovations in Electrical and Electronic Engineering

Proceedings of ICEEE 2021

 Springer

Editors

Saad Mekhilef
Department of Electrical Engineering
University of Malaya
Kuala Lumpur, Malaysia

Margarita Favorskaya
University of Science and Technology
Reshetnev Siberian State
Krasnoyarsk, Russia

R. K. Pandey
Department of Electrical Engineering
Indian Institute of Technology (BHU)
Varanasi, Uttar Pradesh, India

Rabindra Nath Shaw
School of Electrical and Electronics
Engineering
Galgotias University
Greater Noida, Uttar Pradesh, India

ISSN 1876-1100

ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-981-16-0748-6

ISBN 978-981-16-0749-3 (eBook)

<https://doi.org/10.1007/978-981-16-0749-3>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.

The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

The book constitutes selected high-quality papers presented in International Conference on Electrical and Electronics Engineering (ICEEE 2021) jointly organized by School of Electrical, Electronics and Communication Engineering, Galgotias University, India, and School of Engineering, University of Malaya, Malaysia, on January 2–3, 2021. It discusses issues in electrical and electronics engineering and technologies along with AI and IoT applications. An in-depth discussions on various issues under topics provide an interesting compilation for researchers, engineers, and students.

We are thankful to all the authors that have submitted papers for keeping the quality of ICEEE 2021 at high levels. The editors of this book would like to acknowledge all the authors for their contributions and the reviewers. We have received invaluable help from the members of the International Program Committee and the chairs responsible for different aspects of the workshop. We also appreciate the role of Special Sessions Organizers. Thanks to all of them, we had been able to collect many papers on interesting topics, and during the conference, we had very interesting presentations and stimulating discussions.

Our special thanks go to Leopoldo Angrisani, (Editor in Chief, Springer, *Lecture Notes in Electrical Engineering* Series) for the opportunity to organize this guest-edited volume.

We are grateful to Springer, especially to Aninda Bose (Senior Editor, Hard Sciences Publishing), for the excellent collaboration, patience, and help during the evolution of this volume.

We hope that the volume will provide useful information to professors, researchers, and graduated students in the areas of electrical and electronics engineering and technologies along with AI and IoT applications, and all will find this collection of papers inspiring, informative, and useful. We also hope to see you at a future ICEEE event.

Kuala Lumpur, Malaysia
Krasnoyarsk, Russia
Varanasi, India
Greater Noida, India

Saad Mekhilef
Margarita Favorskaya
R. K. Pandey
Rabindra Nath Shaw

ICEEE 2020 Organization

Chief Patrons

Mr. Suneel Galgotia, Chancellor, Galgotias University, India
Mr. Dhruv Galgotia, CEO, Galgotias University, India

Patron

Prof. Bhim Singh, Indian Institute of Technology Delhi, India.
Prof. R. Venkatesh Babu, Pro Vice Chancellor, Galgotias University

General Chairs

Prof. Saad Mekhilef, Dean, School of Engineering, University of Malaya, Malaysia
Prof. B. Mohapatra, Dean, SECE, Galgotias University, India

Conference Chair and Chairman, Oversight Committee

Prof. Rabindra Nath Shaw, Galgotias University, India

Conference Secretary

Prof. Saravanan D, Galgotias University, India

Technical Chairs

Dr. Nishad Mendis, Det Norske Veritas, Australia
Prof. Ankush Ghosh, The Neotia University, India
Prof. Rohit Tripathi, Galgotias University, India

Publication Chairs

Prof. Valentina E. Balas, Aurel Vlaicu University of Arad, Romania
Prof. Sanjoy Das, IGNTU Manipur, India

Springer/ICEEE Liaison

Aninda Bose, Senior Editor, Springer Nature

International Advisory Board

Prof. Vincenzo Piuri, University of Milan, Italy
Prof. Georges Zissis, President, IEEE IAS
Prof. Lakhmi C. Jain, University of Technology, Sydney
Dr. Tamas Ruzsanyi, Ganz-Skoda Electric Ltd., Hungary
Prof. Valentina Balas, University of Arad, Romania
Prof. N. R. Pal, President, IEEE CIS
Prof. T. George, University of Piraeus, Greece
Prof. Yen-Wei Chen, Ritsumeikan University, Japan
Prof. Milan Simic, RMIT University, Australia
Prof. M. Paprzycki, Polish Academy of Sciences
Prof. Maria Virvou, University of Piraeus, Greece
Prof. D. P. Kothari, Ex-Director, IIT Delhi, India
Prof. B. K. Panigrahi, Professor, IIT Delhi, India
Prof. C. Boccaletti, Sapienza University, Italy
Dr. Akshay Kumar, Concordia University, Canada

Conference Report: ICEEE 2021

The 2nd International Conference on Electrical and Electronics Engineering (ICEEE 2021) was jointly organized by School of Engineering, the University of Malaya (UM) (Malay: Universiti Malaya), and Galgotias University, India, on January 2–3, 2021. ICEEE 2021 has emerged as a premier conference to bring researchers, scientists, and industry professionals to discuss novel ideas in the field of computing, power, and communication. This year, the conference was joined by researchers, students, and industry professionals from many countries. Due to the COVID-19 pandemic, the conference was organized online.

In the inaugural session of ICEEE 2021, the opening remarks were given by Shri Dhruv Galgotia, CEO, Galgotias University, India. The guests of honor of the conference were Prof. (Dr.) Pierluigi Siano, Professor, Electrical Energy Engineering University of Salerno, Italy; Prof. (Dr.) Lance C.C. Fung, Professor, Murdoch University, Australia, Director-Elect, IEEE Region 10 (Asia Pacific); Prof. (Dr.) Saifur Rahman, Joseph Loring Professor and Director, Virginia Tech Advanced Research Institute, Virginia, USA; Prof. (Dr.) Saad Mekhilef, Dean, Faculty of Engineering, University of Malaya, Malaysia; Prof. (Dr.) Monica Bianchini, Professor, Information Engineering and Mathematics Department, University of Siena, Italy; Dr. C. P. Ravikumar, Director, Technical Talent Development, Texas Instruments; Prof. (Dr.) Milan Simic, Professor, University Union Nikola Tesla, Belgrade, Serbia; Prof. (Dr.) Hazlie Mokhlis, Professor, Faculty of Engineering, University of Malaya, Malaysia; Prof. (Dr.) Atif Iqbal, Professor, Qatar University, Vice-Chair, IEEE Qatar Section. The vote of thanks was given by the Conference Secretary.

The conference commenced with three keynote speeches (1) ‘New Challenges in VLSI Design and Test’ by Dr. C. P. Ravikumar, Director of Technical Talent Development, Texas Instruments, (2) ‘Research Directions in Electrical Power Distribution System’ by Prof. (Dr.) Hazlie Bin Mokhlis, Faculty of Engineering Professor, University of Malaysia, Malaysia, and (3) ‘Peer-to-Peer Networks—What We Have Been Doing?’ by Prof. (Dr.) Y. N. Singh, Professor, Electrical Engineering, Indian Institute of Technology Kanpur, India. This was followed by concurrent paper presentation panels FD1 to FD2 in two technical sessions. On completion of these two, another two technical sessions FD3 and FD4 started. On the first day, 40 papers were presented with the participation of 91 authors.

The second day began with two parallel technical sessions SD1 and SD2 followed by another two technical sessions SD3 and SD4 with 40 papers in total. At the end of the technical sessions, the following keynote speeches were delivered. (1) ‘Intelligent Transportation Systems’ by Prof. (Dr.) Milan Samic, Professor, University Union Nikola Tesla, Belgrade, Serbia, (2) ‘Grid Integrated Distributed Energy Generation System Standards, Challenges, and Requirements’ by Prof. (Dr.) Atif Iqbal, Professor, College of Engineering, Qatar University, Qatar, (3) ‘Reliable Power Electronics—An Enabler for Renewable Energy Generation and Utilization’ by Prof. (Dr.) Huai Wang Professor, the Faculty of Engineering and Science, Aalborg University, Denmark, and (4) ‘Power Electronics Converters for Solar Power Systems: Current Trends and Challenges’ by Prof. (Dr.) Bhim Singh, Professor, Electrical Engineering Department, Indian Institute of Technology Delhi, India. The keynote sessions were attended by a large number of students and researchers.

ICEEE 2021 concluded with a closing session. Dignitaries present in the closing ceremony were Dr. Saad Mekhilef, Dean, Faculty of Engineering, University of Malaya, Malaysia; Dr. Arun K. Tripathi, Director General, National Institute of Solar Energy, MNRE, Government of India; Dr. Valentina E. Balas, Professor, Aurel Vlaicu University of Arad, Romania; Prof. Bhim Singh, Professor, Indian Institute of Technology Delhi, India; Prof. (Dr.) Milan Simic, Professor, University Union Nikola Tesla, Belgrade, Serbia; Prof. (Dr.) Marcin Paprzycki, Systems Research Institute, Polish Academy of Sciences, Poland. The conference report was given by Dr. Nishad Mendis, Technical Chair for ICEEE 2021. Vote of thanks was given by Dr. B. Mohapatra, Dean, SEECE, Galgotias University, and with this, ICEEE 2021 was concluded.

Rabindra Nath Shaw
Conference Chair
ICEEE 2021

Contents

Lifecycle Assessment of Electricity Generation transition in Ecuador	1
Manuel Ayala-Chauvin, Carlos Samaniego-Ojeda, Genís Riba, and Jorge Maldonado-Correa	
Force and Rotor Displacement Analysis of Bridge Configured Based Induction Machine	11
Sivaramakrishnan Natesan, Gaurav Kumar, and Karuna Kalita	
Smart Grid Cyber vulnerabilities and Mitigation Measures	25
Priya R. Krishnan and Josephkutti Jacob	
Innovative Approach to Harvest Energy from Traffic-Induced Deformation	49
Nilimamayee Samal, O. J. Shiney, and Abhinandan Patra	
Impact Analysis of Single Line to Ground Fault on Industrial Loads Using Typhoon HIL	63
R. Rajesh Kanna, R. Raja Singh, and D. Arun Dominic	
Performance Improvement for PMSG Tidal Power Conversion System with Fuzzy Gain Supervisor Passivity-Based Current Control	81
Youcef Belkhier, Abdelyazid Achour, Rabindra Nath Shaw, and Ankush Ghosh	
Harmonics Minimization in Multilevel Inverter by Continuous Mode ACO Technique	95
Salman Ahmad, Atif Iqbal, Imtiaz Ashraf, and Zahoor Ahmad Ganie	
Loss Allocation Method for Microgrids Having Variable Generation	105
Dibya Bharti	

Contingency Analysis of Power Network with STATCOM and SVC	121
P. S. Vaidya and V. K. Chandrakar	
Comparative Analysis of Modified Firefly Algorithm with Firefly Algorithm MPPT for PV Module Under Partial Shaded Condition	131
S. Sachin, Anusha Vadde, and Sowmyashree	
Active Power Control for Single-Phase Grid Connected Transformerless Inverter Photovoltaic System	145
Janardhan Gurram, G. N. Srinivas, and N. N. V. Surendra Babu	
Data Modeling for Energy Forecasting Using Machine Learning	159
Monika Goyal and Mrinal Pandey	
Adaptive Linear Feedback Energy-Based Backstepping and PID Control Strategy for PMSG Driven by a Grid-Connected Wind Turbine	177
Youcef Belkhier, Abdelyazid Achour, Rabindra Nath Shaw, Walid Sahraoui, and Ankush Ghosh	
Permanent Magnet Synchronous Motor Drive for Electric Vehicle with Efficient Battery Management System	191
Varij Kumar and Saurabh Mishra	
Two-Step Experimental Validation of Impact of Irregular Irradiance on Solar Photovoltaic System's Performance	203
Paresh S. Nasikkar, Chandrakant D. Bhos, and Javed K. Sayyad	
Transfer Learning-Based Novel Fault Classification Technique for Grid-Connected PV Inverter	217
Azra Malik, Ahteshamul Haque, K. V. Satya Bharath, and Zainul Abdin Jaffery	
Improvement of Power System Stability Using FACTS Controller	225
Harsh Mishra, Faisal Raza, and Shagufta Khan	
Performance Evaluation of Solar PV Using Multiple Level Voltage Gain Boost Converter with C-L-C Cell	237
B. M. Kiran Kumar, M. S. Indira, and S. Nagaraja Rao	
Comparative Study for Different Types of MPPT Algorithms Using Direct Control Method	253
Immad Shams, Karam Khairullah Mohammed, Saad Mekhilef, and Kok Soon Tey	
Protection Scheme Based on k-Nearest Neighbour Algorithm for Fault Detection Classification and Section Identification in DC Microgrid	265
Shankarshan Prasad Tiwari, Ebha Koley, and Subhojit Ghosh	

Solar PV-Fed Grid-Interactive PMSM-Driven Water Pumping System with Switched-LC Converter 273
 Rasna Rajan, C. M. Nirmal Mukundan, M. K. P. Muhammad Ramees, and P. Jayaprakash

Impact Analysis of Microgrid in Minimization of Distribution System Losses 283
 Nawaz Ali Warsi, Anwar Shahzad Siddiqui, Sheeraz Kirmani, and Md Sarwar

Performance Estimation and Analysis of 3D Trigate HOI FinFET Using Strained Channel for Reduced Area 289
 Swagat Nanda and Rudra Sankar Dhar

Economic and Environmental Analysis of a Solar-Powered EV Charging System in India—A Case Study 301
 Prashant Shrivastava, TeyKok Soon, Saad Mekhilef, and Furkan Ahmad

Islanding Detection Review Using Intelligence Classifier in Distribution Network 317
 Hairul Nissah Zainudin, Saad Mekhilef, Hazlie Mokhlis, and Safdar Raza

Evaluating Classical and ANN-Based Load Forecasting Techniques Using Univariate and Multivariate Analysis 349
 Kuheli Goswami, Ayandeep Ganguly, Nayan Manna, and Arindam Kumar Sil

Comparative Performance of Different PV Array Topologies Under Partial Shading Condition 367
 Karni P. Palawat, Vinod K. Yadav, Ramjee L. Meena, and Santosh Ghosh

Fault Detection in Power Transmission System Using Reverse Biorthogonal Wavelet 381
 Gaurav Kapoor, Vishesh Kumar Mishra, Rabindra Nath Shaw, and Ankush Ghosh

HWT-DCDI-Based Approach for Fault Identification in Six-Phase Power Transmission Network 395
 Gaurav Kapoor, Pratima Walde, Rabindra Nath Shaw, and Ankush Ghosh

Graphene Properties and Its Utility for High-Frequency Antennas 409
 Subodh Kumar Tripathi and Rohit Tripathi

Comparative Analysis of MOSFET, FINFET and GAAFET Devices Using Different Substrate and Gate Oxide Materials 417
 Ritik Koul, Mukul Yadav, and Rajeshwari Pandey

VLSI Design of a Split Parallel Two-Dimensional HEVC Transform 431
 Ainy Haziyah Awab, Ab Al-Hadi Ab Rahman, Izam Kamisian, and Mohd Shahrizal Rusli

Convergence of Novel Iterative Learning Control Methods for a Class of Linear Discrete-Time Switched Systems	441
D. R. Sahu and Nitish Kumar Singh	
A Scalable VLSI Architecture for Illumination-Invariant Heterogeneous Face Recognition	459
Swagata Bhattacharya, Somsubhra Talapatra, Debotosh Bhattacharjee, and Amlan Chakrabarti	
Enhanced Performances of SAC-OCDMA System Operating with Different Codes	473
Walid Sahraoui, Hakim Aoudia, Angela Amphawan, Smail Berrah, Youcef Belkhier, and Rabindra Nath Shaw	
Phasor Estimation of Power Signals in a Smart Grid Environment Using Sigmoid Adaptive Filter	487
Soumili Saha, Prateek Bose, and Sarita Nanda	
Design and Analysis of Brain-Implantable Antenna for Neural Signal Transmission	495
Vipan Kumar Gupta and Payal Mehra	
A Novel Design of FSS-Based Absorber Integrated Microstrip Antenna	505
Priyanka Das and Kaushik Mandal	
Off-State Leakage Concern in Scaling Nanowire FETs	513
Rajiv Ranjan Thakur, Nitin Chaturvedi, and Nidhi Chaturvedi	
Performance Analysis of FBMC and OFDM with MIMO for Wireless Communications	521
Yedukondalu Kamatham and Sushmitha Pollamoni	
Magnetization Pattern Study of Unit Domain Multiferroic Nanomagnet for Spintronics Devices	533
Amlan Sen, Rabindra Nath Shaw, and Ankush Ghosh	
Advanced RSSI-Based Wi-Fi Access Point Localization Using Smartphone	543
Debajyoti Biswas, Suvankar Barai, and Buddhadeb Sau	
A Novel Miniaturized Fractal Antenna on Circular Patch with Defected Ground Structure (DGS)	555
Sakshi Singh, Jitesh Kumar, Arvind Gaur, and Rahul Vishnoi	
Rotating Acoustic Reflector Parameter Trade-Off for Near-Outdoor Audio Event Detection	567
Ganesh Bhagwat, Sangeeth Jayaprakash, and Anirban Bhowmick	

Performance Evaluation and Synthesis of FIR Filters Using Various Multipliers Algorithms 583
 Aniket Kumar and R. P. Agarwal

Design and Implementation of Bluetooth Low Energy Link Layer Controller Using Dataflow Programming 593
 Nur Atiqah Abdul Latib, Ab Al-Hadi Ab Rahman, and Mohd Shahrizal Rusli

Bandwidth and Gain Enhancement of Circular MPA Using Twin-Layer Stacked Antenna 605
 Harshit Srivastava, Usha Tiwari, and Pallavie Tyagi

Rooftop Antenna for Vehicular Application 617
 Muskan Bhatia, Mansi Dimri, and Brajlata Chauhan

Design of Compact Quad Band Monopole Antenna Using Open-Ended and Meandered Slot 627
 Raj Kumar and Yedukondalu Kamatham

Design of a HNL Index Guiding Silica PCF for SCG: Applications in Dental Optical Coherence Tomography System 635
 Mohammad Kamrozzaman Kiron, Md. Masum Hossen, Nguyen Hoang Hai, and Md. Anwar Hossain

A Machine Learning Approach for the Land-Type Classification 647
 Arvind Kumar and Tejalal Choudhary

A Tiny CNN Architecture for Medical Face Mask Detection for Resource-Constrained Endpoints 657
 Puranjay Mohan, Aditya Jyoti Paul, and Abhay Chirania

Analysis of Darknet Traffic for Criminal Activities Detection Using TF-IDF and Light Gradient Boosted Machine Learning Algorithm 671
 Romil Rawat, Vinod Mahor, Sachin Chirgaiya, Rabindra Nath Shaw, and Ankush Ghosh

Risk Detection in Wireless Body Sensor Networks for Health Monitoring Using Hybrid Deep Learning 683
 Anand Singh Rajawat, Kanishk Barhanpurkar, Rabindra Nath Shaw, and Ankush Ghosh

AI and IOT-Based Model for Photovoltaic Power Generation 697
 Harshit Kumar Huneria, Pavan Yadav, Rabindra Nath Shaw, D. Saravanan, and Ankush Ghosh

Combating Challenges in the Construction Industry with Blockchain Technology 707
 Priyanka Singh, Debarati Sammanit, Paritosh Krishnan, Krishna Mohan Agarwal, Rabindra Nath Shaw, and Ankush Ghosh

Dark Web—Onion Hidden Service Discovery and Crawling for Profiling Morphing, Unstructured Crime and Vulnerabilities Prediction 717
 Romil Rawat, Anand Singh Rajawat, Vinod Mahor, Rabindra Nath Shaw, and Ankush Ghosh

Suspicious Big Text Data Analysis for Prediction—On Darkweb User Activity Using Computational Intelligence Model 735
 Anand Singh Rajawat, Romil Rawat, Vinod Mahor, Rabindra Nath Shaw, and Ankush Ghosh

Development of Prediction Models to Determine Compressive Strength and Workability of Sustainable Concrete with ANN 753
 Priyanka Singh, Saurav Bhardwaj, Saurav Dixit, Rabindra Nath Shaw, and Ankush Ghosh

Smart Attendance and Progress Management System 771
 Lathika Krishnapillai, Sangavi Veluppillai, Adonija Akilan, V. Naomi Saumika, K. P. Dhammika H. De Silva, and M. P. A. W. Gamage

Performance and Parametric Analysis of IoT’s Motes with Different Network Topologies 787
 Rajkumar Gaur and Shiva Prakash

Fruit Classification Using Deep Learning 807
 Jinu Lilly Joseph, Veena A. Kumar, and Santhosh P. Mathew

Orthonormal Bayesian Convolutional Neural Network for Detection of the Novel Coronavirus-19 819
 Romin Chetankumar Jariwala and Madhusudana Rao Nalluri

Classification of Chest Diseases from X-ray Images on the CheXpert Dataset 837
 Hasan Nabeel Saleem, Usman Ullah Sheikh, and Saifulnizam Abd. Khalid

Feature-Based Overview of Online Comments of Web-Based Healthcare Products 851
 Saroj Kushwah, Bharti Kalra, and Sanjoy Das

A Mutual Authentication and Key Agreement Protocol for Vehicle to Grid Technology 863
 Azmi Iqbal, Akber Ali Khan, Vinod Kumar, and Musheer Ahmad

Fire Detection and Real Time Monitoring Systems Through IoT Sensors 877
 Sudip Suklabaidya and Indrani Das

Intelligent Physical Access Control System Through Three-Stage Verification Using IoT 889
 Rohit Tripathi, Manoj K. Shukla, and Yogesh Kumar

Descriptive Indic Answer Script Evaluation Using Deep Learning 901
Deepak Senapati, Abinash Panda, Siddhant Mund,
and Sunil Kumar Mohapatra

Blockchain Aided Predictive Time Series Analysis in Supply Chain System 913
G. A. Dhanush, Kiran S. Raj, and Priyanka Kumar

Regional Language Code-Switching for Natural Language Understanding and Intelligent Digital Assistants 927
S. B. Rajeshwari and Jagadish S. Kallimani

Dimensionality Reduction for Face Recognition Using Principal Component Analysis Based Big Bang–Big Crunch Optimization Algorithm 949
Supreet Grewal and C. Rama Krishna

OntoVidRec: A Staged Knowledge Aggregation Scheme for Annotations-Based Video Retrieval Using Ontology Matching 957
Gerard Deepak and A. Santhanavijayan

An Approach for Retrieval of Text Documents by Hybridizing Structural Topic Modeling and Pointwise Mutual Information 969
K. Vishal, Gerard Deepak, and A. Santhanavijayan

Deep Belief Convolutional Neural Network with Artificial Image Creation by GANs Based Diagnosis of Pneumonia in Radiological Samples of the Pectoralis Major 979
Tathagat Banerjee, Dhruv Batta, Aditya Jain, S. Karthikeyan,
Himanshu Mehndiratta, and K. Hari Kishan

About the Editors

Saad Mekhilef is IET Fellow and IEEE Senior Member. He is Associate Editor of IEEE Transaction on Power Electronics and Journal of Power Electronics. He is Professor at the Department of Electrical Engineering, University of Malaya, since June 1999. He is currently Dean of the Faculty of Engineering and Director of Power Electronics and Renewable Energy Research Laboratory, PEARL. He is the author and the co-author of more than 400 publications in international journals and proceedings (253 ISI journal papers) and five books with more than 17,000 citations and 64 H-index; 116 Ph.D. and master students have graduated under his supervision. He has six patents. He was frequently invited to give keynote lectures at international conferences. Prof. Saad listed by Thomson Reuters (Clarivate Analytics) as one of the Highly Cited (Hi Ci) engineering researchers in the world and included in the Thomson Reuters' The World's Most Influential Scientific Minds: 2018. He is actively involved in industrial consultancy for major corporations in the power electronics projects. His research interests include power conversion techniques, control of power converters, renewable energy and energy efficiency.

Margarita Favorskaya is Professor and Head of the Department of ICT at Reshetnev Siberian State University of Science and Technology, Russian Federation. Professor Favorskaya is Member of KES organization since 2010, IPC Member and Chair of invited sessions of over 30 international conferences. She serves in many international journals as Reviewer, Guest Editor and Book Editor (Springer). She is the author or the co-author of 200 publications and 20 educational manuals in Computer Science and Electrical Engineering. She co-authored/co-edited books for Springer recently. She supervised nine Ph.D. candidates and presently supervising four Ph.D. students. Her main research interests are in the allocations of artificial intelligence and information technologies in Electrical Engineering, Electronic Engineering, Computer Engineering, Bio-Medical Engineering, Communication and related disciplines.

R. K. Pandey is a known power system expert who is not only Professor of Electrical Power Systems at IIT (BHU) and executed Project on “Design and Development of Smart Energy Grid Architecture with Energy Storage” at BESCOM Bengaluru

with Smart Metering Operational Control Room. He has also established An Intelligent Solar Power Control Architecture for 6.2 MWp BHU Solar Plant. He has been on deputation from July 2016 to August 2020 to the Ministry of Power Government of India as Director General, National Power Training Institute.

Rabindra Nath Shaw is Senior Member of IEEE (USA), currently holding the post of Director, International Relations, Galgotias University India. He is an alumnus of the applied physics department, University of Calcutta, India. He has more than eleven years of teaching experience in leading institutes like Motilal Nehru National Institute of Technology Allahabad, India, Jadavpur University and others in UG and PG levels. He has successfully organized more than fifteen international conferences as Conference Chair, Publication Chair and Editor. He has published more than fifty Scopus/WoS/ISI indexed research papers in international journals and conference proceedings. He is Editor of several Springer and Elsevier books. His primary area of research is optimization algorithms and machine learning techniques for power system, IoT application, renewable energy, and power electronics converters. He also worked as University Examination Coordinator, University MOOC's Coordinator, University Conference Coordinator and Faculty In-Charge, Centre of Excellence for Power Engineering and Clean Energy Integration.

Lifecycle Assessment of Electricity Generation transition in Ecuador



Manuel Ayala-Chauvin , Carlos Samaniego-Ojeda , Genís Riba ,
and Jorge Maldonado-Correa 

Abstract Ecuador's energy mix has greatly reduced its dependency on fossil fuels the last 15 years, down to a marginal role (5%) in electricity generation in 2017. The development plan for the Ecuadorian power network aims to keep adding hydropower to meet the increasing demand. A prospective lifecycle assessment (LCA) of the future power network (2012–2050) can determine the feasibility of the development plan and its environmental sustainability in the long run. For a quantitative analysis of the energy transition over the entire lifecycle, the simulation software® Global Emission Model of Integrated System (GEMIS) is used. The results show that the current development path of the Ecuadorian energy system reduces the emissions of CO₂ per kWh generated by 65% due to the large share of renewable energies, mainly hydropower, which costs 1% of Gross Domestic Product. The obtained LCA footprints are similar to the literature benchmarks.

Keywords Energy transition · Sustainability · Renewable energy · Ecuador · Electricity generation · GEMIS

M. Ayala-Chauvin (✉)

SISAU research Group, Facultad de Ingeniería y Tecnologías de la Información y Comunicación, Universidad Tecnológica Indoamérica, Ambato 180103, Ecuador
e-mail: mayala@uti.edu.ec

C. Samaniego-Ojeda · J. Maldonado-Correa

Facultad de la Energía, las Industrias y los Recursos Naturales No Renovables, Universidad Nacional de Loja, Av Pio Jaramillo, Loja EC110103, Ecuador
e-mail: carlos.samaniego@unl.edu.ec

J. Maldonado-Correa

e-mail: jorge.maldonado@unl.edu.ec

G. Riba

Centro de Diseño de Equipos Industriales, Universidad Politécnica de Cataluña, Barcelona 08028, España
e-mail: genis.riba@cdei.upc.edu

1 Introduction

1.1 Background

The need for an energy transition is indisputable according to [1], in which the author crosses proven and possible energy resources data with the forecasted evolution of global energy needs in a medium and long term. According [2], climate change, fossil fuels depletion and the economic vulnerability open up the debate to change the energy model and encourage the study of transition dynamics from a historical point of view. The renewable energy transition entails the expansion and integration of sustainable renewable energy generation systems into the existing infrastructures, replacing existing systems with the new ones in a progressive, orderly and rapid manner [3].

According to the Ecuadorian Ministry of Electricity and Renewable Energy (MEER), the National Interconnected System (SNI) had an effective installed capacity of 4415.3 MW in 2009, 2032 MW of which of hydroelectric capacity and 2286 MW of thermal capacity, to cover a demand of more than 2000 MW. This may suggest an excessive power reserve. However, thermal plants require fuel and hydroelectric plants require certain water reserves, which not always are available. The dual systems complemented each other: In times of drought, the cheaper hydropower was insufficient, and the thermal park, besides being expensive, did not have enough capacity to meet energy demand on its own [4].

During the last decade, Ecuador carried out an ambitious process to strengthen and transform its energy matrix. The MEER invested approximately M\$ 4600 in various renewable energy generation projects. Mazar hydroelectric came online in 2010, Villonaco wind farm started operation in 2013 [5], Paute Integral came online in 2015 and Coca Codo Sinclair in 2016. As new hydrogeneration came online, older thermal plants were progressively removed from the grid. All these projects helped bolstering Ecuador's energy independence and turned Ecuador into a net energy exporter. In merely 10 years, Ecuador completed an energy transition from fossil fuels to renewable energy in the electricity sector [6].

Comparing the energy mix on 2006 and 2017, Ecuador went from a 36% dependency on thermal generation and 18% on imports to being self-sufficient with a 93% renewable generation (Fig. 1).

Energy efficiency measures and the promotion and integration of renewable energy sources (RE) are two of the main pillars of the energy concept of Ecuador. Ecuador is committed to reduce the total primary energy supply (TPES), reduce greenhouse gas (GHG) emissions and reduce energy consumption in 2020.

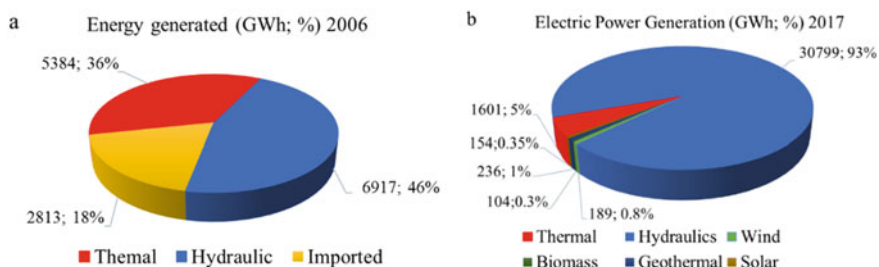


Fig. 1 Generation of the electricity subsector 2006 (a) and 2017 (b)

1.2 Related Works

Scientific literature includes a number of LCA studies about of the different power generation projects implemented in Ecuador. Below we briefly describe some of them:

The article presented in [7] is a recent study that aims to analyze the environmental impacts associated with variations in the participation of hydroelectricity in the energy mix in Ecuador between 2012 and 2050 using lifecycle methods. The results show that increasing the share of hydropower in the energy mix considerably reduces the environmental impacts of electricity generation per kWh. In addition, the projected Global Warming Potential (GWP) per kWh of the electricity in 2050 remains below 2016 levels for the scenarios where the hydroelectric capacity is implemented according to the national electrification plan.

The authors in [8] present a study in which they use climate change scenarios for an optimization model of Ecuador's energy system through 2050, in order to examine the lower-cost options for the Ecuadorian electricity system dominated by hydroelectric energy. The results show that, in the long term, hydroelectric energy will continue to be one of the most profitable and low-emission technologies in the Ecuadorian electricity sector. In addition, they mention that, according to the analysis based on models, changes in water availability could induce a variation in the generation of hydroelectric power to supply 29–86% of the total electricity demand. Another relevant aspect of the conclusion of the authors of this study is that the current policy of the Ecuadorian government is based on the assumption that there will only be small changes in future hydrological conditions and the levels of runoff available to power hydroelectric projects.

A forecasting model for the Ecuadorian energy sector, using the LEAP model, is presented in [9]. This study analyzes the behavior of the energy matrix based on energy forecasting and efficiency policy scenarios, applying a bottom-up analysis and considering the latest policies in Ecuador. The authors conclude that the model and the proposed considerations will result in a final energy consumption of 158 million Barrels of Oil Equivalent (BOE) by 2030, in which the transport sector is the main energy consumer. They also indicate that there will be a 10-year period of surplus of electricity generation; the maximum surplus is 60% over five years, which

will be reduced to 19% in 2030. The main cause of this reduction is the increase in household electricity consumption, due to the introduction of electrical appliances [9].

The Electrification Master Plan (2012–2050) [10], issued by the MEER, presents five development paths for the Ecuadorian power grid. It is based on power generation and other government projects, and accounts for all generation connected to the National Interconnected System (SNI), but it does not include self-generators (e.g., oil fields). Demand projections consider the National Program for Efficient Cooking, which aims to replace liquefied petroleum gas (LPG) with electricity in domestic efficient cooking and hot water appliances with state funding and subsidies. Induction cookers are expected to use 5150 GWh by the end of the substitution process. Electricity demand projections include 596.9 GWh/year on savings due to efficiency programs, and 3794 GWh/year increased demand due to increasing activity in mining, cement, steel, oil and Yachay (City of Knowledge).

Regarding mobility, projections consider an extra 351 GWh/year due to the incorporation of electric vehicles for light transport and public transport (Quito subway and Cuenca tram). Finally, the Pacific Refinery in Aromo near the city of Manta in Manabí province will require around 375 MW for the refining facility, other production processes and the development of the area. Although the likelihood of this scenario is high due to the availability of technical studies, project implementation largely depends on political will.

In this study, a prospective lifecycle assessment (LCA) of different power generation projects implemented in Ecuador was performed. Finally, the results of the environmental analysis of power generation and operation, sustainability indicators of economic and maintenance (O&M) costs are presented. The remainder of this paper is structured as follows: Sect. 2 describes the method; Sect. 3 presents the results obtained, and finally, Sect. 4 presents the conclusions.

2 Method

The Global Emission Model for Integrated Systems (GEMIS) is a European lifecycle analysis database program provided by International Institute for Sustainability Analysis and Strategy (IINAS) (Fig. 2).

GEMIS gives a detailed description of all the process steps of an energy system and the calculation of the primary energy consumption involved in the process, the emissions and the mass and energy flows [11]. We use this model to calculate lifecycle emissions of power plants [12–16].

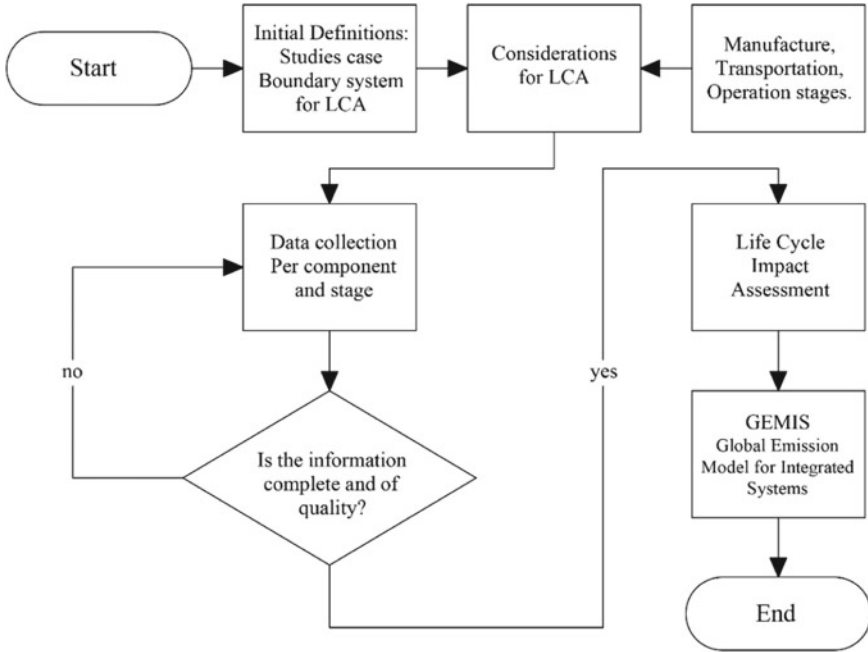


Fig. 2 Flowchart of adopted method. Source developed by the authors

3 Results

The results of the simulation can be summarized into four categories: the projected demand by the end of the period, the costs to build and operate the power generation to meet the demand and the resulting air pollution and sustainability indicators.

3.1 Projected Demand

See Tables 1 and 2.

Table 2 shows the results of the simulation performed with GEMIS yields a projected demand of 114,232 GWh in 2050. This result is very close to that obtained by [10] in the average growth scenario (hypothesis 5), in which the projected demand amounts to 114,223 GWh.

Table 1 Power plant data introduced into GEMIS. *Source* developed by the authors

GEMIS central	Power (MW)	Units	Total power (MW)	Operation (h)	Energy GEMIS (GWh/year)
Hydro-dam-big-1100 MW	1100	9	9900	4585	45,392
Hydro-dam-big-generic	200	12	2400	5477	13,145
Hydro-dam-medium-generic	50	101	5050	6035	30,481
Bagasse-ST-CO-2010	20	5	100	2964	296.4
Gas-GT-big-generic diesel	50	13	650	972	631.8
Gas-CC-big-generic	200	4	800	6299	5,039
Oil-heavy-ST-big-generic	410	1	410	5050	2,071
Diesel motor-big-generic	50	20	1000	4740	4,740
Oil-heavy-ST-small-generic	33.24	1	33.24	6240	207.42
Diesel motor-big-generic	113.1	1	113.18	769	87.04
Hydro-dam-CO-2010	650.0	500	3,250	360	1,170
Solar-PV-utility-CN	20	18	360	2222	800
Windfarm-big-generic	10	169	1,690	1735	2,932.15
Biogas-dieselmotor-generic	10.7	1	10.7	6000	64.2
Geothermal-ST-big-generic	100	9	900	7973	7175.7
Total			26,667		114,232

Table 2 Demand and installed capacity. *Source* developed by the authors

Year	Demand (GWh)	Installed capacity (MW)	Maximum power (MW)
2012	19,316.26	5135.89	3207.00
2050	114,232.06	26,667.00	18,966.00

3.2 Total Costs

Total costs include capital investments for the construction of power plants and fixed and variable O&M costs. Fixed O&M costs refer to those that do not vary significantly with the level of generation: general and administrative expenses, routine and preventive maintenance of the facilities, maintenance of structures and land and others. Variable O&M costs refer to those that vary with generation, e.g., fuel [17–19] (Tables 3 and 4).

Total costs amount to 1% GDP (2012–2050), which is US\$ 6,308,171 million. Although the analysis is not comprehensive (without updating prices, inflation, etc.), the results are a good reference.

Table 3 O&M costs

Electric power plant	Total electricity production (GWh/year)	O&M (USD/MWh)	Total cost 2012–2050 (M USD)
Hydroelectric up to 1000 MW	43,623.79	3.86	8725
Hydroelectric more than 1000 MW	45,393	3.86	19,850
Biomass TV bagasse	296.36	42.64	480
Biomass MCI biogas	64.2	32.33	86
Solar PV	800	0.5% capital cost-year	1632
Eolic	2932.98	17	3366
Geothermal	7176	35	10,375
TV (Fuel oil)	2070.57	84.9	6680
TV (Diesel)	207.43	111.31	877
TG (Diesel)	631.71	169.41	4067
TG (GN)	0	126.81	409
MCI (Fuel oil)	4740.16	72.28	13,321
MCI (Diesel)	87.02	97.25	322
Combined cycle	5038.92	97.35	18,970
Total	113,062.14		89,161

Table 4 Greenhouse gases

Year	CO ₂ Eq	Carbon dioxide, CO ₂	Methane, CH ₄	Nitrous oxide, N ₂ O
2012	7062.258	6634.052	13.088	229
2050	15,175.149	13,810.044	42.692	406

3.3 Air Emissions

Global warming is the gradual increase in the temperature of the Earth as a result of the increase in the concentration of greenhouse gases in the atmosphere. The global warming potential is determined based on the CO₂ equivalent: 100 year greenhouse warming potential (GWP) of methane is 25 and that of N₂O is 298 [18–20].

Pollution results provided by GEMIS with an energy demand of 114232 GWh are the following:

CO₂-eq emissions are calculated considering the full lifecycle. In general terms, the following lifecycles were considered: 50 years for large dam hydroelectric plants, 20 years for solar and wind power plants and 15 years for thermal power plants.

Table 5 Sustainability indicators

Data and indicators	2012	2050
Population	15,520,973	27,777,933
GDP M USD (2010)	78,945	290,810
Power Consumption, GWh	19,316	114,233
CO ₂ -eq emissions, ton	7,062,258	15,175,149
external dependence	4%	1%
renewable share	63%	88%
MWh/hab	1.24	4.11
MWh/M USD	244.68	392.81
tCO ₂ /MWh	0.37	0.13
tCO ₂ /M USD	89.46	52.18
tCO ₂ /hab	0.46	0.55

3.4 Sustainability Indicators

The following table shows the main sustainability indicators that allow comparisons between 2012 and 2050:

Table 5 shows the main sustainability indicators. It is important to note that external dependence is minimal, 1%.

In addition, it is important to note a renewable share in 2050 over 80%, which is reflected in the low emissions per MWh generated.

4 Conclusions

LCA studies can be used to forecast the future environmental impacts of a power network. This article explores the sustainability of the power network growth plan envisioned by the Ecuadorian government through 2050.

Global Emission Model for Integrated Systems (GEMIS) was used to determine the environmental performance of electricity generation. The obtained LCA footprints are compared to the literature data and resulting in similar.

The indicators show that external dependence is minimal in 2050 (1%), as this scenario considers the operation of Pacific Refinery, avoiding the import of diesel for thermal power plants. The production of electrical energy from fossil sources (diesel, fuel oil and NG) is marginal, but necessary for demand peaks and to improve the robustness of the electrical system.

The participation of renewable energies is relevant since it is close to 90%, which is reflected in the low emissions per MWh generated, which is a 65% reduction compared to 2012 emissions. Hydroelectric energy is the main primary energy source, with 78% of share, due to its great potential for expansion, since the rivers that

flow into the Pacific Ocean are complemented by those flowing into the Amazon basin, minimizing the effects of droughts. Per capita electricity consumption is higher in 2050 than in 2012, indicating a greater development due to the impact of the industrialization planned by the government. The current model implanted in the Ecuadorian energy matrix costs 1% of Gross Domestic Product.

Energy transition in Ecuador follows the path of sustainable development, with appropriate technologies, socially acceptable, and economically profitable.

Acknowledgements This research was supported by Secretaría Nacional de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT), Ecuador; SISA research Group, Facultad de Ingeniería y Tecnologías de la Información y Comunicación, Universidad Tecnológica Indoamérica, Ambato, Ecuador; Universidad Nacional de Loja, Ecuador; and Centro de Diseño de Equipos Industriales of Universidad Politécnica de Cataluña, BarcelonaTech, Spain.

References

1. R. Carles, *Energy Resources and Crisis the End of 200 Unrepeatable Years*, 2nd edn. (Oficina de Publicacions Acadèmiques Digitals de la UPC, Barcelona, 2011).
2. F. Kern, K.S. Rogge, The pace of governed energy transitions: agency, international dynamics and the global Paris agreement accelerating decarbonisation processes? *Energy Res. Soc. Sci.* **22**, 13–17 (2016). <https://doi.org/10.1016/j.erss.2016.08.016>
3. E. Kyritsis, J. Andersson, A. Serletis, Electricity prices, large-scale renewable integration, and policy implications. *Energy Policy* **101**, 550–560 (2017). <https://doi.org/10.1016/j.enpol.2016.11.014>
4. L. Vintimilla, “*Apagones y soluciones*” *El sector energético: presente y perspectivas*. CIE Corporación para la Investigación Energética (2002)
5. M. Ayala, J. Maldonado, E. Paccha, C. Riba, Wind power resource assessment in complex terrain: villonaco case-study using computational fluid dynamics analysis. *Energy Procedia* **107**, 41–48 (2017). <https://doi.org/10.1016/j.egypro.2016.12.127>
6. “Plan Maestro de Electrificación 2013–2022 VOL III,” 2013.
7. A.D. Ramirez, A. Boero, B. Rivela, A.M. Melendres, S. Espinoza, D.A. Salas, Life cycle methods to analyze the environmental sustainability of electricity generation in Ecuador: Is decarbonization the right path? *Renew. Sustain. Energy Rev.*, **134**, 110373. <https://doi.org/10.1016/j.rser.2020.110373>
8. P.E. Carvajal, F.G.N. Li, R. Soria, J. Cronin, G. Anandarajah, Y. Mulugetta, Large hydropower, decarbonisation and climate change uncertainty: modelling power sector pathways for Ecuador. *Energy Strateg. Rev.* **23**, 86–99 (2019). <https://doi.org/10.1016/j.esr.2018.12.008>
9. P. L. Castro Verdezoto, J. A. Vidoza, W.L.R. Gallo, Analysis and projection of energy consumption in Ecuador: energy efficiency policies in the transportation sector. *Energy Policy* **134**, 110948 (2019). <https://doi.org/10.1016/j.enpol.2019.110948>
10. “Plan Maestro de Electrificación 2013–2022 VOL II,” (2013)
11. B. Guezuraga, R. Zauner, W. Pölz, Life cycle assessment of two different 2 MW class wind turbines. *Renew. Energy* **37**(1), 37–44 (2012). <https://doi.org/10.1016/j.renene.2011.05.008>
12. R. Madlener, S. Stagl, Sustainability-guided promotion of renewable electricity generation. *Ecol. Econ.* **53**(2), 147–167 (2005). <https://doi.org/10.1016/j.ecolecon.2004.12.016>
13. A. Makhlof, T. Serradj, H. Cheniti, Life cycle impact assessment of ammonia production in Algeria: a comparison with previous studies. *Environ. Impact Assess. Rev.* **50**, 35–41 (2015). <https://doi.org/10.1016/j.eiar.2014.08.003>

14. H. Yang, L. Ma, Z. Li, A method for analyzing energy-related carbon emissions and the structural changes: a case study of China from 2005 to 2015. *Energies* **13**(8), 1–24 (2020). <https://doi.org/10.3390/en13082076>
15. N. Thonemann, A. Schulte, D. Maga, How to conduct prospective life cycle assessment for emerging technologies? A systematic review and methodological guidance. *Sustainability* **12**(3) (2020). <https://doi.org/10.3390/su12031192>
16. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
17. R.N. Shaw, P. Walde, A. Ghosh, A new model to enhance the power and performances of 4×4 PV arrays with puzzle shade dispersion. *Int. J. Innovat. Technol. Exp. Eng.* **8**(12) (2019)
18. V. Rakotoson, J.-P. Praene, A life cycle assessment approach to the electricity generation of French overseas territories. *J. Clean. Prod.* **168**, 755–763 (2017). <https://doi.org/10.1016/j.jclepro.2017.09.055>
19. Updated capital cost estimates for utility scale electricity generating plants (2013)
20. IPCC, *Climate Change 2007-The Physical Science Basis* (United Kingdom and New York: Cambridge University Press, 2007)

Force and Rotor Displacement Analysis of Bridge Configured Based Induction Machine



Sivaramakrishnan Natesan, Gaurav Kumar, and Karuna Kalita

Abstract Bridge configuration winding (BCW) is one of the stator winding method to generate a controllable transverse force to attenuate the unbalanced magnetic pull (UMP) in the induction machine. This paper investigates the analysis of radial and tangential forces which are acting on the air gap of a bridge configured based induction machine by using FEM. An Opera 2D finite element solver has been used to measure the radial and tangential force distribution around the mid-air gap of the BCW based induction machine. The rotor of the FE BCW induction machine model has been shifted in the *X-axis* direction for 60% eccentricity with the stator to analyze the bridge currents. The forces were calculated and compared for the following cases: (i) when $1 V_{dc}$ is supplied at the three-phase bridge points, (ii) without any external supply. A controller has been developed and tries to make an attempt for the active vibration control in the BCW induction machine. The rotor displacements were measured for same input conditions as given in the FE model and presented in the frequency domain.

Keywords Bridge configured winding · Unbalanced magnetic pull · Vibration control

1 Introduction

Non-circular profile of the rotor due to the manufacturing defect and an imperfect alignment of rotor with the stator in the induction machine can produce a transverse force called unbalanced magnetic pull (UMP) toward the shortest air gap. This UMP can be increased along with the rotor eccentricity when the speed of the rotor

S. Natesan (✉)

Saveetha Engineering College, Chennai, Tamil Nadu, India

G. Kumar

Indian Institute of Information Technology, Bhagalpur, Bihar, India

K. Kalita

Indian Institute of Technology, Guwahati, Assam, India

increases. This effect would lead to the imbalances of the magnetic field around the air gap. Further, this results in the continuous contact between the rotor and stator, vibration, and noises which may worsen the performances of the machine. Belman et al. [1] developed an analytical air gap permeance model for calculation of magnetic flux density, and they used total magnetic energy to derive the UMP in two-pole induction machine. UMP can be controlled by using various types of stator winding. A group of research people are working on the mitigation of UMP by using various kind of stator winding connections. There are some significant transformations in the existing work on controlling of UMP by providing series winding, parallel winding, double set of winding, and single set of winding connection. The research group of Arkkio [2, 3] and Chiba [4] have exploited the capability of electrical motors and generators to exert controllable transverse forces on the rotor and to generate bearing forces for bearing-less application. Both of these groups have achieved this by deploying double sets of stator winding, one set of winding is used to develop the torque on the rotor, and the second set is used to generate the transverse force. Khoo [5] invented a new winding called bridge configured winding which can be able to generate torque as well as transverse force with the help of primary stator winding itself. The advantage of the BCW can provide three different voltage or current supply sources which are isolated from each other. The BCW primary winding configuration is advantageous in the sense that the same coils are used for the torque production and transverse force generation, whereas in dual set of winding scheme, the stator slots accommodate two different winding connection which effects the torque producing capability of the machine. The characteristics of induction machines are mostly analyzed by finite element method. Over 100 years, the research people uses FEM for the simulation of electrical machines. The existing works includes the complex magneto-dynamic formulation in the presence of single harmonic frequency in both stator and slip of the rotor. It also includes the incorporation of rotor movement, effect of end-windings, and skewness of rotor bar. The time-stepping method is quite useful in the sense that it can be used to calculate all the parameters of the machine. The skewed rotor effect and homopolar flux are not possible to incorporate in a 2D FE model. Tenhunen and Arkkio [6] used 2D multi-slice model to incorporate skewed rotor effect. However, it comes at the cost of computational effort. Also, the homopolar flux has not been considered for the finite element analysis [7]. In addition to these formulations, there are few formulations in which the transient dynamic behavior of the induction machines is analyzed and reported by Vassent et al., [8] and later by Ho et al. [9]. This transient dynamic analysis is accomplished by coupling the mechanical motion equation and the coupled field-circuit equation. Dorrell [10] studied and reviewed the effect of asymmetric magnetic field in the air gap on UMP. An experimental analysis is performed to study the effect of tooth saturation. The effect of slotting has not been presented because the effect of slotting comes in terms of audible noise and high frequency vibration not the UMP. Parallel paths reduces UMP, whether winding is on the rotor or the stator. Dorrell and Smith, [11, 12] have developed an analytical model for the calculation and measurement of UMP and later verified it experimentally. The model was developed for motors having series and parallel winding connections

along with rotor eccentricities. A series-connected winding forces the individual coil currents to be constant and hence provides a sinusoidal current density distribution in the machine. A fully parallel-connected winding on the other hand allows the coil currents to redistribute themselves and forces a constant voltage across each coil. Debortoli et al. [13] demonstrated the effect of parallel winding on UMP production. In the parallel winding, the stator circuits which is located at the farthest point from the rotor has lower inductance. Hence, a higher amount of current flows in those circuits, and it controls UMP even if the rotor is eccentric. These currents are called as circulating current. Laiho et al. [14] developed analytical model for an electromechanical rotor with built-in force actuator has been proposed. The concept of unbalanced field distribution has been utilized for the force generation in active magnetic bearing. Arkkio [15] calculated the forces by using time-stepping method in finite element analysis of the magnetic field. The author has done the analysis in 2D domain for the field in the magnetic core. The Crank–Nicholson method is used for the modeling the time-dependent fields. The simulation results are compared for rotor broken bar, broken end-ring, eccentricity, and number of rotor slots. The largest forces were produced by broken end-rings. It has been noticed that the circulating currents and the core saturation have reduced the magnitude of UMP. Kalita and Laiho [16] have calculated analytically the lateral force in the air gap in bridge configured induction motor. Natesan et al. [17] have developed an experimental rig which is based on bridge configured induction machine to detect the unbalance present in the machine by analyzing the frequency components of bridge currents. Natesan et al. [18] investigated bridge currents in an asymmetric rotor with the stator FE model. The FE model has 10% static eccentricity of the air gap for the analysis. Gaurav et al. [19] demonstrated the vibration control by incorporating the additional external bridge supply of a bridge configured induction motor. Gaurav et al. [20] include the effect of various eccentricity conditions and the bridge currents in the FE model to analyze the controlling of vibration.

The main objective of the present work is that to demonstrate the radial and tangential forces which are acting on the mid-air gap of a bridge configured induction machine by using a finite element model. The advantageous and working principle of bridge configured winding has been explained in [18]. The FE model has 60% static eccentricity of the rotor with the air gap, i.e., the rotor is moved 0.75 mm along the *X-axis*. The induction machine has a radial air gap length of 1.25 mm. There are many literature works that have been published on the demonstration of bridge currents. However, there are only few literature works that have been presented on the radial and tangential forces acting on the air gap of an induction machine. Therefore, this work is mainly focused on the following issues:

- Initially, the radial and tangential forces were calculated for three cases: (i) without BCW connection, (ii) with BCW connection, and (iii) when 1 Vac source supplied at bridge points of 60% static eccentric BCW induction machine FE model.
- A controller rig has developed in a 37 kW, three-phase, and four-pole induction machine. 1 V_{dc} has been supplied at the three-phase bridge points as an external supply source.

- The rotor displacements and bridge currents were measured for three different cases: (i) main bridge OFF and ON conditions, (ii) controller bridge ON condition, and (iii) an amplitude of 1 V_{dc} as external supply at three-phase controller bridge points.

2 FE Analysis of BCW Based Induction Machine Using Opera 2D Package

A detailed explanation of FE modeling of an induction machine, incorporation of bridge configured winding scheme, and machine parameter have been presented in [18]. The rotor of the FE BCW induction machine model has been shifted in the X -axis direction for 60% eccentricity with the stator to analyze the bridge currents. The rotation motion analysis has been utilized for the force analysis. Figure 1 shows the 2D FE model of a bridge configured based induction by using Opera 2D FE solver. The FE model has been simulated for the various following conditions: (i) Without BCW connection, (ii) with BCW connection, and (iii) when 1 Vac source is supplied at the three-phase bridge points.

2.1 Results and Discussions of FE Analysis

With reference to circuit connection shown in [18], the switches $S1$, $S2$, and $S3$ are called as bridge points in the FE induction model. The bridge points $S1$, $S2$, and $S3$ are not short-circuited in open condition which is referred as without bridge connection, and the bridge points $S1$, $S2$, and $S3$ are short-circuited which is referred as with bridge connection. The induction machine FE model has simulated with the main supply voltage frequency of 20 Hz. The three-phase bridge currents were measured, and only phase A bridge currents have been plotted in frequency domain. However, it is expecting that the phase B and phase C bridge currents are also having

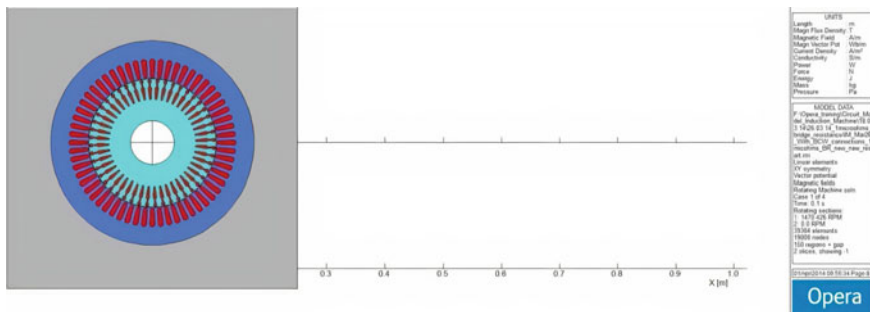


Fig. 1 2D FE model of a bridge configure based induction machine by using Opera 2D FE solver

the similar frequency components as same as like phase A bridge currents. Figure 2 shows the phase A bridge currents in frequency domain without an external supply for the main supply frequency of 20 Hz. Figure 3 shows the phase A bridge currents in frequency domain when 1 V_{ac} supply has given at the ridge points for the main supply frequency of 20 Hz.

The dominant frequency components of the bridge currents for dynamic eccentricity condition are given by Eq. (1) as presented by Gaurav et al. [19].

$$f_{\text{bridge}} = f_{\text{supply}} \pm f_{\text{rotor}} \tag{1}$$

where f_{supply} is the frequency of main supply and f_{rotor} is the frequency of rotor. It has been observed in Fig. 1 that the frequencies of the bridge current are 10.05 and

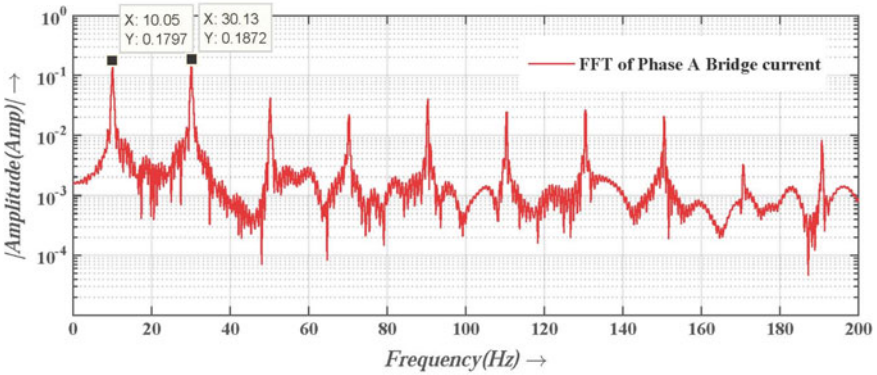


Fig. 2 Phase A bridge currents in frequency domain without an external supply for the main supply frequency of 20 Hz

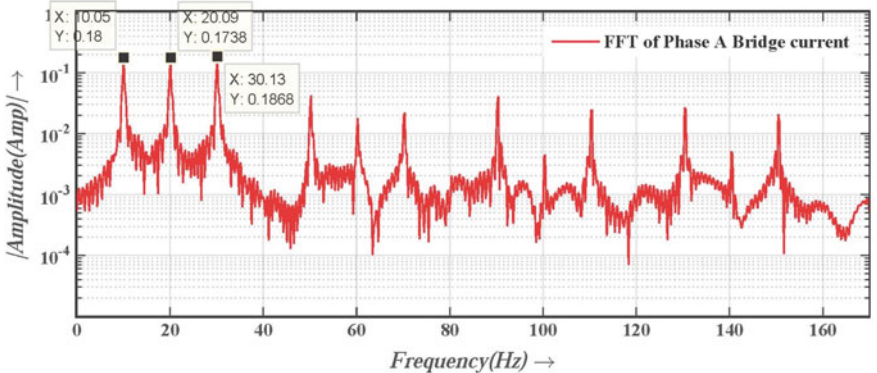
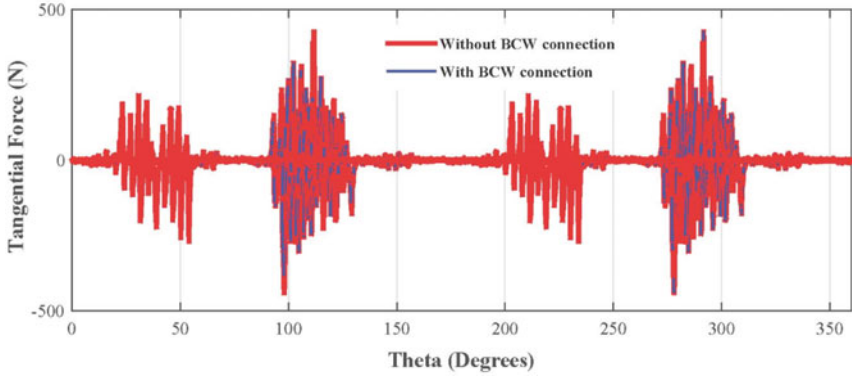


Fig. 3 Phase A bridge currents in frequency domain when 1 V_{ac} supply has given at the ridge points for the main supply frequency of 20 Hz

Table 1 Comparisons of frequency components of bridge currents with the eccentricity case

FE model	Frequency components (Hz)	Eccentricity case
No external supply	10.05, 30.13 ($f_{\text{supply}} \pm f_{\text{rotor}}$)	Dynamic eccentricity
1 Vac external supply	10.05, 20, 30.13 ($f_{\text{supply}}, f_{\text{supply}} \pm f_{\text{rotor}}$)	Mixed eccentricity

**Fig. 4** Comparison of tangential forces acting on the mid-air gap of with and without BCW connection of 60% eccentricity induction machine FE model

30.13 Hz, i.e., ($f_{\text{supply}} \pm f_{\text{rotor}}$) and represents Eq. 1 that the FE model has dynamic eccentricity condition. Similarly, it has also been noticed in Fig. 2 that the frequencies of the bridge current are 10.05, 20, and 30.13 Hz, i.e., ($f_{\text{supply}}, f_{\text{supply}} \pm f_{\text{rotor}}$) and represents the FE model that has mixed eccentricity condition. Table 1 summarizes the comparisons of frequency components of bridge currents with the eccentricity case.

The radial and tangential forces have been calculated by using Maxwell force distribution by integrating the flux density in the horizontal and vertical direction over the mid-air gap region of the induction machine. Equation (2) and Eq. (3) show the tangential and radial force component acting over the mid-air gap region.

$$F_{\text{Tangential}} = \int_0^{2\pi} \frac{b_a^2(\theta, t)}{2\mu_o} \sin(\theta) r l d\theta \quad (2)$$

$$F_{\text{radial}} = \int_0^{2\pi} \frac{b_a^2(\theta, t)}{2\mu_o} \cos(\theta) r l d\theta \quad (3)$$

where b_a is magnetic flux density, θ is angle of arc, r is the mid-air gap radius = 109.875 mm, and l is the length of the rotor core = 212 mm.

Figures 4 and 5 show the comparison of tangential and radial forces acting on the mid-air gap of with and without BCW connection of 60% eccentricity induction machine FE model, respectively. It has seen that the tangential and radial force components are greatly reduced when the bridge is being connected. The tangential forces are balanced with each other over the zero axis line which is the reason for the rotation of the rotor. It has also been observed that the radial forces are unbalanced, i.e., acting only in one direction. This clearly shows that the radial forces are the responsible for the deviation of rotation of the rotor with the stator. This is happening due to the 60% of the dynamic eccentricity present in the BCW induction FE model.

Figures 6 and 7 show the comparison of tangential and radial forces acting on the mid-air gap of with BCW connection and when 1 V_{ac} supplied at all three-phase bridge points of 60% eccentricity induction machine FE model, respectively. It has

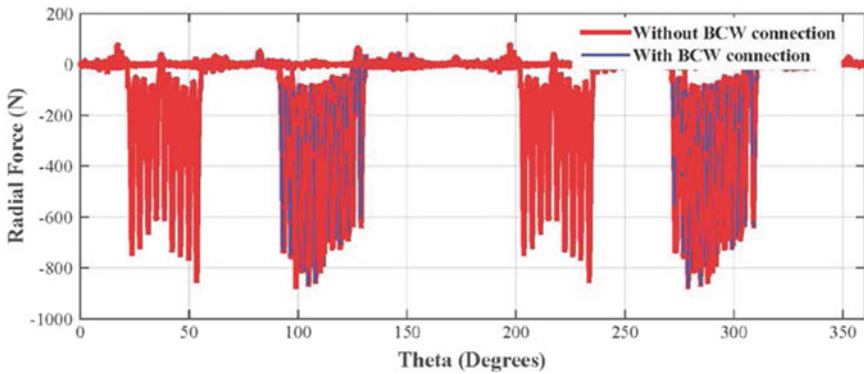


Fig. 5 Comparison of radial forces acting on the mid-air gap of with and without BCW connection of 60% eccentricity induction machine FE model

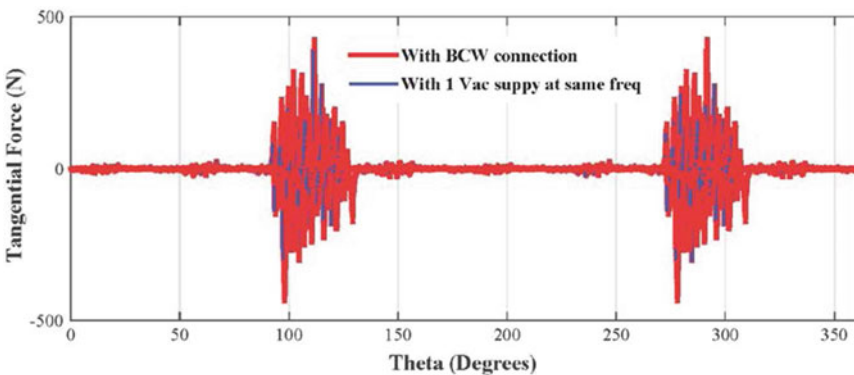


Fig. 6 Comparison of tangential forces acting on the mid-air gap of BCW induction machine with no external supply and 1 V_{ac} supplied at all three-phase bridge points of 60% eccentricity induction machine FE model

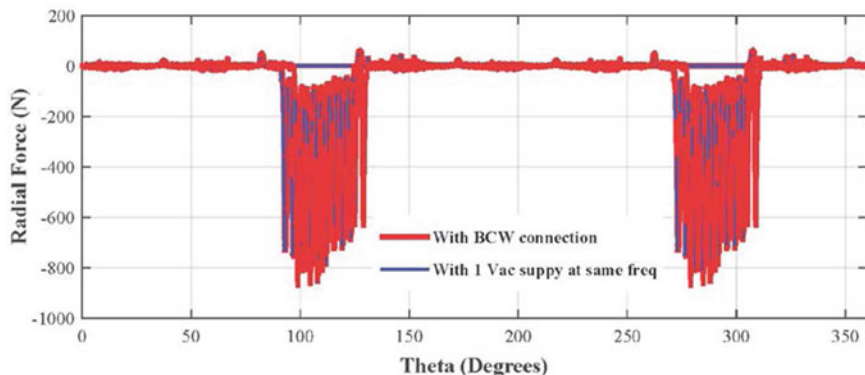


Fig. 7 Comparison of radial forces acting on the mid-air gap of BCW induction machine with no external supply and $1 V_{ac}$ supplied at all three-phase bridge points of 60% eccentricity induction machine FE model

seen that there is a small amount of reduction in the radial forces though the amplitude $1 V_{ac}$ may not be high enough and further study is required.

3 Development of Controller Rig Set Up

The development of the experimental setup and the incorporation of the bridge configured winding in a four-pole, three-phase, 37 kW induction machine have been illustrated in [18]. A controller has been developed to incorporate the three-phase external voltage supply at the bridge points. The present section illustrates the development of the controller which have been utilized for the active vibration control in a bridge configured based induction machine. The controller has different components mainly isolation amplifier, Copley controller, isolation transformer, and a rectifier with filter. Table 2 shows the name, make, and model of the devices which are used in the development of controller.

Three isolation amplifiers have been used to break the loops between common ground level of reference signal and the controller for safety purpose. A $15 V_{dc}$

Table 2 Devices used in controller

S. No.	Name of the device	Make of the device	Model
1	Isolation amplifier	Analog devices	AD202JN
2	Servo amplifier	Copley controller	5221 CE
3	Isolation transformer	RKD	RKD 100/2 × 24
4	Current transducer	LEM transducer	LTS 15-NP
5	Voltage transducer	LEM transducer	LV 25-P

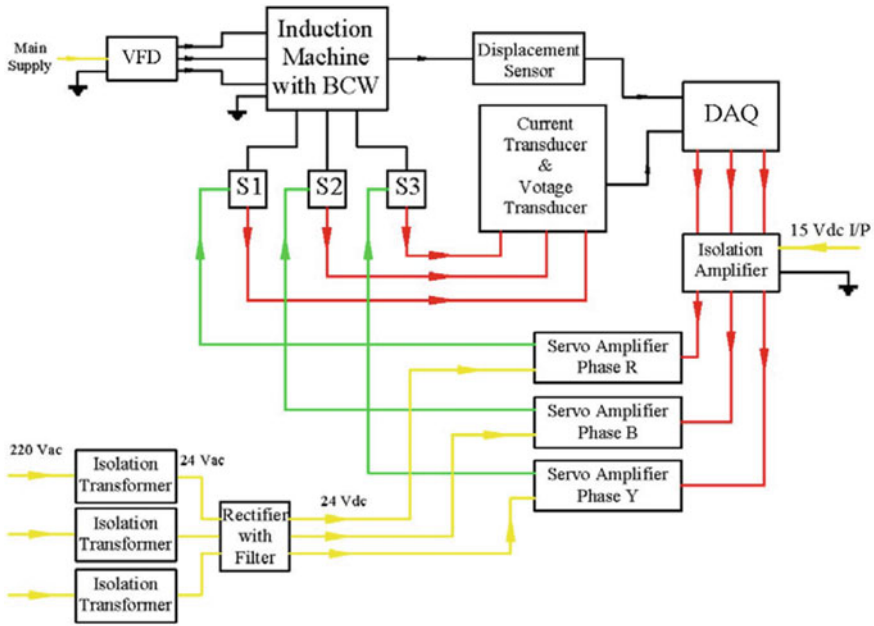


Fig. 8 Flow diagram of connection of devices

battery power source has been used as a power supply for the isolation amplifier. Three DC servo amplifiers have been used to supply the required controlled voltage in three bridges as per the reference signal provided by the DAQ system. Figure 8 shows the connection flow diagram of the components used in the controller. Three isolation transformers have been used to step down the 220 V_{ac} to 24 V_{ac} supply. A rectifier filter circuit has been used to convert 24 V_{ac} signal to a 24 V_{dc} signal, and this 24 V_{dc} supply has been used as a power source for the servo amplifiers.

3.1 Experimental Analysis

The three-phase controller bridge currents and corresponding rotor displacements (in X and Y directions) have been measured for the main supply frequency of 20 Hz. The three-phase bridge currents and rotor orbits have been obtained for two cases, (i) bridge OFF and (ii) bridge ON (controller). The switches for the controller bridge connections are shown in Fig. 9. With reference to Fig. 8, controller bridge OFF condition means that the switches S1, S2, and S3 are in open condition. Controller bridge ON condition means that the switches S1, S2, and S3 are in closed condition. The bridge currents were measured at bridge points S1, S2, and S3 as soon as the controller bridge points are short-circuited. Figures 10 and 11 show the FFT of the rotor responses in X- and Y-directions, respectively, at 20 Hz supply frequency

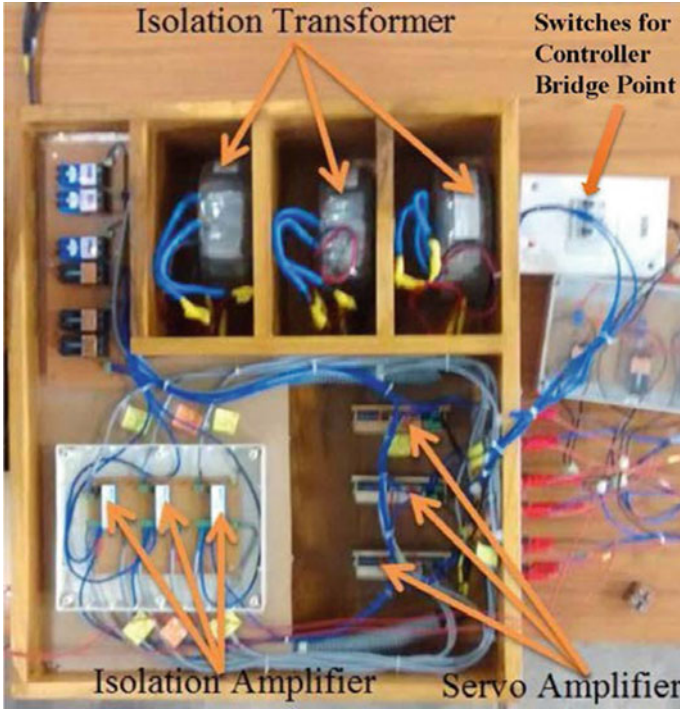


Fig. 9 Controller setup for active vibration control

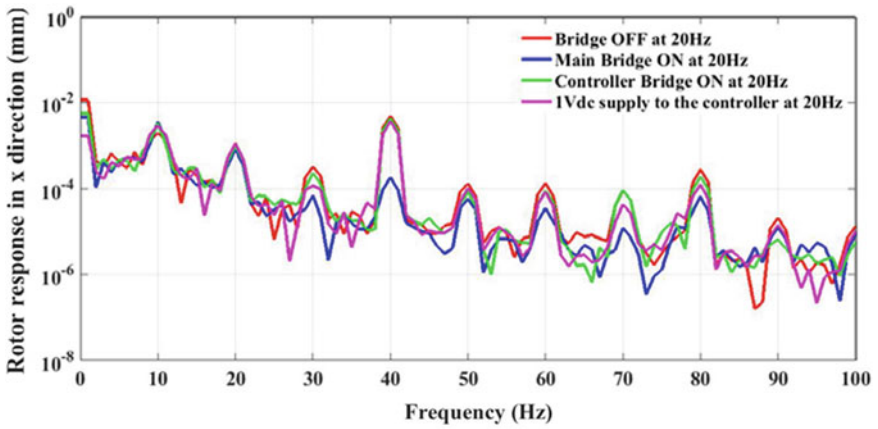


Fig. 10 Comparison of FFT of the rotor responses in X-direction for main bridge OFF, ON, controller bridge on, and 1 V_{dc} supply conditions for a main supply frequency of 20 Hz

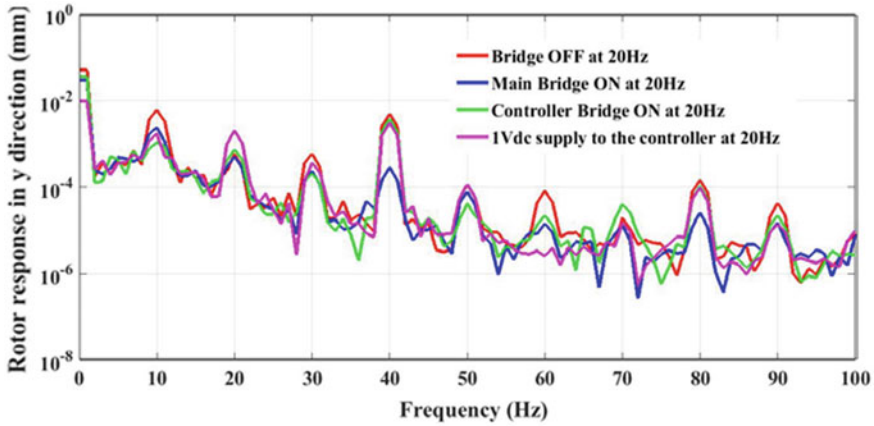


Fig. 11 Comparison of FFT of the rotor responses in Y-direction for main bridge OFF, ON, controller bridge on and 1 V_{dc} supply conditions for a main supply frequency of 20 Hz

in bridge OFF and controller bridge ON conditions. It has been observed that the frequency component of the rotor responses along X- and Y- directions for a main supply frequency of 20 Hz are 0 Hz ($f_{supply} - f_{supply}$), 10 Hz ($f_{supply} - 1/2 f_{supply}$), 20 Hz (f_{supply}), 30 Hz ($f_{supply} + 1/2 f_{supply}$), 40 Hz ($2 f_{supply}$), and 50 Hz ($f_{supply} + 3/2 f_{supply}$), where f_{supply} is the frequency of the main supply. The double supply frequency component 40 Hz ($2 f_{supply}$) of rotor responses which have been greatly reduced for main bridge ON (passive vibration) condition when compared with the bridge OFF condition which is not great as for Controller bridge ON condition as well as with controller external supply conditions. This may happen due to the additional frequency components 40 Hz ($2 f_{supply}$) and 60 Hz ($3 f_{supply}$) which are present in the controller bridge currents which are not present in the main bridge currents.

Figure 12 shows the comparison of FFT of bridge currents for main bridge OFF, ON, controller bridge ON, and 1 Vac supply conditions. The frequency component of the bridge currents are 10 Hz ($f_{supply} - f_{rotor}$), 20 Hz (f_{supply}), and 30 Hz ($f_{supply} + f_{rotor}$), where f_s is the frequency of the main supply and f_{rotor} is the rotor frequency. Since the machine has been run at no-load condition, $f_{rotor} = 1/2 f_{supply}$. Figure 13 shows the comparison of rotor orbit for main bridge OFF, ON, controller bridge ON, and 1 V_{ac} supply conditions. It has been clearly visible that the amplitude of rotor orbit has been reduced and tries to move toward the concentric position with the stator.

4 Conclusions

Numerical simulations have been carried out to calculate the radial and tangential forces which present in the mid-air gap region of the BCW induction machine

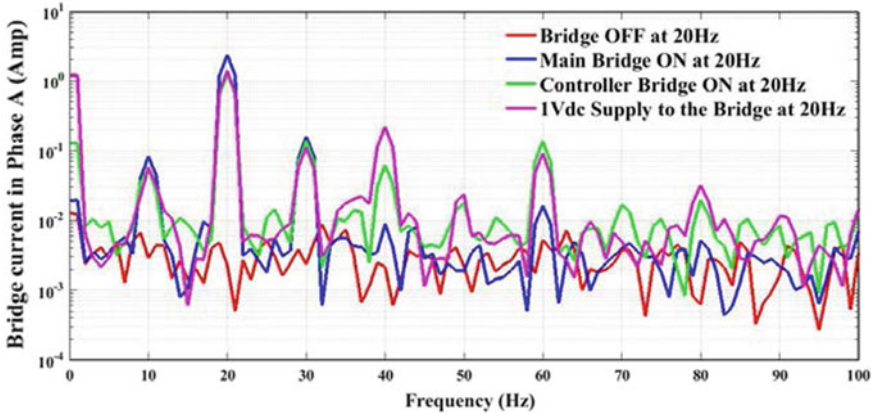


Fig. 12 Comparison of FFT of bridge currents for main bridge OFF, ON, controller bridge ON, and 1 V_{dc} supply conditions for a main supply frequency of 20 Hz

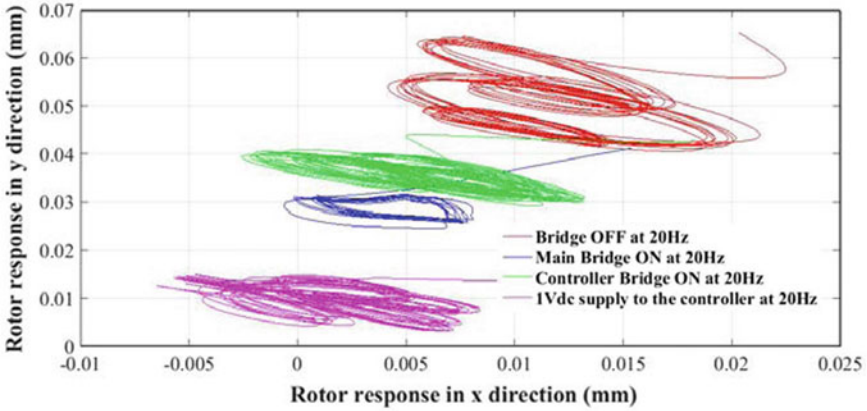


Fig. 13 Comparison of rotor orbit for main bridge OFF, ON, controller bridge ON, and 1 V_{dc} supply conditions for a main supply frequency of 20 Hz

FE model. The radial and tangential forces are compared and presented in order to demonstrate the effect of bridge configured winding in an induction machine. The development of the controller for active vibration control has been presented. The rotor responses and bridge currents are obtained for main supply frequency of 20 Hz for bridge OFF, main bridge ON, and controller bridge ON conditions. The frequency components of the bridge currents in controller bridge ON condition are $(f_{\text{supply}} - f_{\text{rotor}})$, (f_{supply}) and $(f_{\text{supply}} + f_{\text{rotor}})$, where f_s is the frequency of the main supply and f_{rotor} is the rotor frequency. The results obtained from the experiments confirmed that the machine has mixed eccentricity. The rotor responses confirmed that the main bridge ON condition suppresses the $2 f_{\text{supply}}$ component

of the UMP. Controller bridge ON is unable to suppress the $2f_{\text{supply}}$ component of the rotor response. However, there is an additional frequency component of the bridge currents in controller bridge ON condition and that components are $2f_{\text{supply}}$ and $3f_{\text{supply}}$. This perhaps introduce by the controller dynamics and needs further investigation. The bridge ON condition reduces the static eccentricity by shifting the rotor toward the stator center. There are challenges for active vibration control. When the switches of the bridge controller are short-circuited, it changes the nature and magnitude of the current flowing across the bridges compared to the current flowing across the bridges when bridge points S1, S2, and S3 were short circuited directly as in case of passive control.

References

1. R. Belmans, A. Vandenput, W. Geysen, L. Herverlee, Calculation of the flux density and the unbalanced magnetic pull in two pole induction machines, *archivfur Elektrotechnik* **70**, 151–161 (1987)
2. A. Sinervo, A. Arkkio, Eccentricity related forces in two-pole induction motor with four-pole stator damper winding analyzed using measured rotor orbits. *IEEE Trans. Magnet.* **49**(6), 3029–3037 (2013)
3. A. Laiho, K. Tammi, J. Orivuori, A. Sinervo, K. Zenger, A. Arkkio, Electromechanical interaction in eccentric-rotor cage induction machine equipped with a self-bearing force actuator. *J. Syst. Des. Dynam.* **3**(4), 519–529 (2009)
4. A. Chiba, D.T. Power, M.A. Rahman, Characteristics of a bearing less induction motor. *IEEE Trans. Magnet.* **3** (4), 5199–5201 (1991)
5. W.K.S. Khoo, Bridge configured winding for poly phase self-bearing machines. *IEEE Trans. Magnet.* **41**(4), 1289–1295 (2005)
6. A. Tenhunen, A. Arkkio, Modelling of induction machines with skewed rotor slots. *IEE Proc. Electr. Power Appl.* **148**(1), 45–50 (2005)
7. P.P. Silvester, R.L. Ferrari, *Finite Elements for Electrical Engineers* (Cambridge University Press, 1996)
8. E. Vassent, G. Meunier, A. Foggia, G. Reyne, Simulation of induction machine operation using a step by step finite element method coupled with circuits and mechanical equations. *IEEE Trans. Energy Convers.* **14**(3), 304–309 (1999).
9. L. Ho, W. Fu, H.C. Wong, A novel approach to circuit-field-torque coupled time stepping finite element modeling of electric machines. *IEEE Trans. Magnet.* **36** (4), 1886–1889 (2000)
10. D.G. Dorrell, Experimental behaviour of unbalanced magnetic pull in 3-phase induction motors with eccentric rotors and the relationship with tooth saturation. *IEEE Trans. Energy Convers.* **14**(3), 304–309 (2000)
11. A.C. Smith, D.G. Dorrell, Calculation and measurement of unbalanced magnetic pull in cage induction motors with eccentric rotors. Part 1: Analytical model. *IEE Proc.* **143**, 93–201 (1996)
12. A.C. Smith, D.G. Dorrell, Calculation and measurement of unbalanced magnetic pull in cage induction motors with eccentric rotors. Part 2: experimental investigation. *IEE Proc.* **143**, 202–210 (1996)
13. M.J. DeBortoli, S.J. Salon, D.W. Burow, C.J. Slavik, Effects of rotor eccentricity and parallel windings on induction machine behavior: a study using finite element analysis. *IEEE Trans. Magnet.* **29**(2), 1676–1682 (1993)
14. A. Laiho, K. Tammi, K. Zenger, A model-based flexural rotor vibration control in cage induction electrical machines by a built-in force actuator. *Electr. Eng.* 407–421 (2008)

15. A. Arkkio, Calculation of radial forces in cage induction motors at start-the effect of rotor differential. IEEE Conf. Publ. **444**, 36–40 (1997)
16. K. Kalita, A. Laiho, Dynamics of bridged-configured built in force actuator for vibration control in 4-pole cage induction machine, in *National Symposium on Rotor Dynamics, NSRD2011, Chennai* (2011)
17. S. Natesan, G. Kumar, K. Kalita, M. Rahman, Unbalance detection in flexible rotor using bridge configured winding based induction motor, in *Proceedings of the 1st International and 16th National Conference on Machines and Mechanisms (iNaCoMM2013)* (2013)
18. S. Natesan, K. Kalita, V. Samala, Numerical and experimental investigation of bridge currents of an induction machine equipped with bridge configured winding. *Int. J. Electr. Comput. Eng. (IJECE)*, **144**, 1262–1274 (2015)
19. G. Kumar, K. Kalita, K. Tammi, Vibration control using BCW induction motor. *Procedia Eng.* **5**(6), 94–101 (2016)
20. G. Kumar, K. Kalita, K. Tammi, Analysis of bridge currents and UMP of an induction machine with bridge configured winding Using coupled field and circuit modeling. *IEEE Trans. Magnet.* **54**(9), 1–16 (2018)

Smart Grid Cybervulnerabilities and Mitigation Measures



Priya R. Krishnan and Josephkutti Jacob

Abstract Cybersecurity defend and safeguard the computing assets including programs in a cybernetwork thereby providing reliability, integrity, and confidentiality of data transfer. In this paper, a detailed description of the communication framework, a functional block, network types, communication functionalities, reliability, and security aspects are provided. Elaborate elucidation of cybersecurity standards and cyberattack mitigation techniques is reviewed. Current research trends in the field of cybersecurity and future scopes are analyzed.

Keywords Security · Communication · Vulnerability

1 Introduction

Cybersecurity is the practice of securing and defending the electric power generators, substations, transmission and distribution lines, controllers, smart meters, collector nodes, distribution and transmission control centers, and programs from digital attacks. Information attack can be in the form of malware, virus, trojans, spyware, ransomware, adware, and botnets. Unsecured Wi-Fi networks, vulnerabilities in the software, lack of awareness among staff, and negligence are responsible for data theft and hacking. These malicious activities are carried out to attain business, financial, political/intellectual details, sensitive personal data, and customer database and of mere curiosity. Numerous research papers have reported about the prevailing trends in cyberattack, their consequences, and mitigation measures. Hackers can infiltrate and alter the data in a smart grid driven by modern computer technologies and information management systems. Because of the lack of stiff regulatory norms, weakness in IEC 62351 protected smart grid control systems, and loopholes in IEC 61850 protocols, the smart grid has become a primary target for attackers. Dagle in [1] has given the fundamental aspects of cybersecurity in the electric power grid. Issues, drift, and opportunities in the power grid sector are presented outstandingly. Kaster

P. R. Krishnan (✉) · J. Jacob

Electrical and Electronics Department, Cochin University of Science and Technology, Kochi, India
e-mail: priyarkrishnan@cusat.ac.in

and Sen in [2] give an idea about cybersecurity threats. They emphasized the need for industry collaborated research works and elaborated on the technical challenges in this area. Jarmakiewicz et al. in [3] dig up to scratch and provided outstanding and virtuous ideas in the field of cybersecurity. The detailed procedure for the design and implementation of the power grid protection system is presented. Rawat et al. in [4] highlighted the concept of information attack based on the network type, namely home area network (HAN) and wide area network (WAN). They also furnish the influence of each attack on system integrity. Shapsough et al. in [5] elaborated on the information challenges which can cause repercussions on customer privacy. Current security solutions in the network management and protocols are also dealt with. Liang et al. in [6] made a detailed description of grid infrastructure and its security concerns and discuss the areas in the security domain where studies have to be carried out. The concept of the Internet of energy is the main focus in the work of Dari et al. [7]. They gave impotence to specialize intrusion detection system (IDS) and reliability concerns. Wang and Lu in [8] identified the underline differences between the power grid and other networks. Cryptographic and corrective measures for information security are also well explained in [9] the study regarding the accuracy of placement of equipment in the power grid which is mentioned. The authors emphasized the need for a design tool that automates the examination of each link. Sun et al. in [10] designed an IDS and anomaly detection system. Hardware in a loop and cyber-physical system (CPS) testbed have been used to access the power grid security methods. JIANG in [11] proposed a fault location technology that utilizes smart meters and remote fault indicators. A new model based on mixed-integer linear programming is proposed for finding fault location. Adishesu et al. in [12] introduced a scheme for conflict resolution by adding resolve filters. In [13–15], detection and resolution methods of firewall policies are dealt with. Firm and reliable mode bus substation protocols have been proposed in [16–19]. Evaluation of the defenselessness of the SCADA system [19], PLCs [20], and susceptibility of power entities to information attack needs special attention. Vulnerability assessment of SCADA and EMS has been reported in [19, 21]. Theft of data from AMI infrastructure and fault data injection is alarmingly increasing [22–25].

North American Electric Reliability Cooperation (NERC) [26] provides an information technology framework for the dependable and definite operation of the smart grid. National Institute of Standards and Technology (NIST) [27, 28] published NISTIR 7628 for reliable operation of the power grid. A detailed description of the IEC 6180 protocol analysis is carried out in [29–33]. In [29], Khalid et.al. studied the importance of IEC 61850 in supporting smart metering communication. In [30], Elbaset et.al. have discussed the generic object-oriented substation event [GOOSE] function in IEC 61850 protocol. They highlighted the gravity of the IEC 61850 protocol in substation automation. A model that permits the communication time assessment of [GOOSE] messages is proposed in [31]. In [32], Mattos et.al used a graphical interface based on a packet generator that permits the users to create and sent packet according to GOOSE and manufacturing message specification (MMS). Mekkanen and Kauhaniemi in [33] have shown the implementation of low-cost intelligent electronic devices based on IEC 61850–7-420 protocol. In [34], Rajkumar et al.

injected spoofed GOOSE data frames into the communication network of substation for doing vulnerability analysis. In [35], Sathyadevan et al. enumerate the advantages, disadvantages, and outlook of a power system integrated with the 61,850 protocol.

In the remaining of this paper, Sect. 2 explains the smart grid framework, Sect. 3 describes the different network types, and Sect. 4 describes the communication functionalities to be fulfilled in a smart grid. Communication for message transfer is given in Sect. 5 and standards in Sect. 6. Cybersecurity aspects are given in Sect. 7 and smart grid architectural model in Sect. 8. Sections 9 and 10 explains cybersecurity solution threats and standards, respectively. Current research trends and future scope are given in Sect. 12 and finally conclusions in Sect. 13.

2 Fundamental requirements in a smart grid communication system include the functional blocks as shown in Fig. 1

Smart grid framework includes seven domains which are listed below.

2.1 Bulk Generation Domain

It includes all energy providing resources such as hydroelectric power plants, nuclear power plants, coal-fired power plants, diesel-fired power plants, geothermal power plants, solar, wind power plants, and ESS. An important need is to store electricity as energy storage system (ESS) due to the intermittent nature of renewable energy sources (RES). Bulk generation domain requires communication to interact with the transmission domain [36]. Generation companies need a communication network to interact with the load side, distributed energy resources (DER), PEVs, AMIs, and sensors. It also needs to communicate with the operation domain.

2.2 The Transmission Domain

It is used for managing demand and supply. A regional transmission operator or independent system operator (ISO) manages the activities of this domain. It provides an interface between generation and distribution. Information is captured from the grid and sent to control centers through the communication channel. Transformers and sensors send data to the EMS of transmission systems where it is analyzed. Control centers need to communicate with substation devices. The bidirectional control response is scrutinized by the transmission domain.

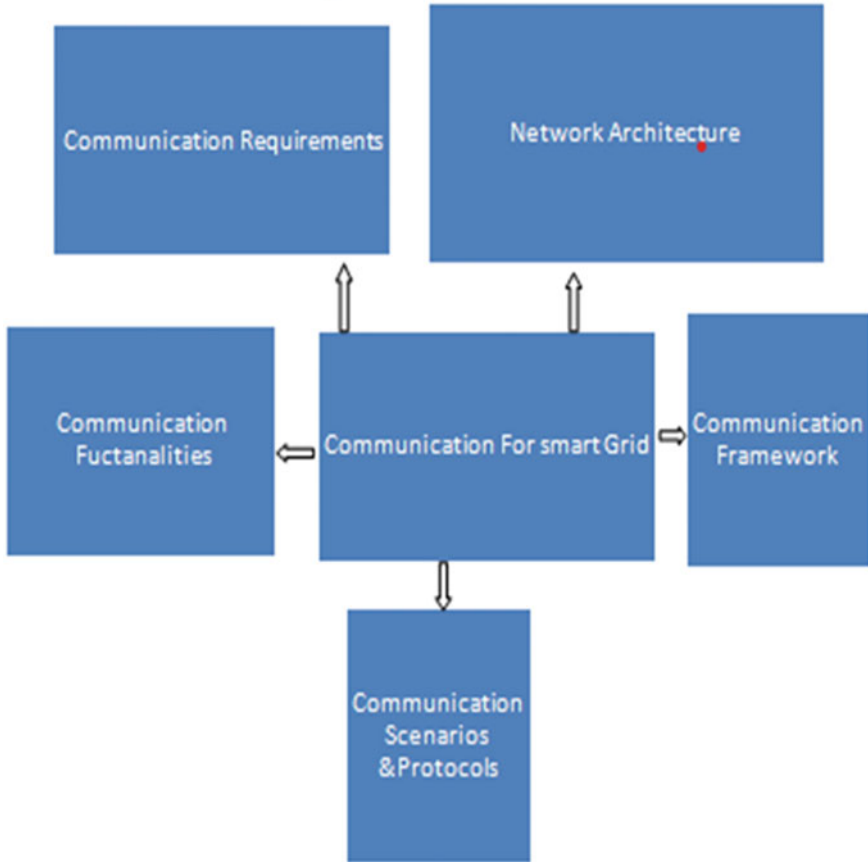


Fig. 1 Fundamental requirements in a smart grid communication system

2.3 Distribution Domain

It includes transformers and feeders and is managed by a distribution management system. The main function of this domain is to deliver electricity with reliability and quality. It communicates with transmission operators to check how much power is available from the transmission system and conduct a discussion with the customer domain to analyze the selling of power. Distributed energy resources (DER), plug-in electric vehicles (PEV), advanced metering infrastructures (AMI), and sensors are also connected in this domain. Bidirectional communication is needed in this domain.

2.4 Operation Domain

The main function of the operation domain is to dispatch electricity and to control the optimization of the EMS and distribution management system (DMS.). Security issues, voltage regulation issues, and frequency variation issues are handled by the operation domain. Communication is carried out with generation, transmission, and distribution domains. Information regarding fault management, planning, and maintenance is collected by this domain. Articulation with SCADA helps in load flow analysis, state flow analysis, and restoration analysis.

2.5 Market Domain

Market domain frames profitable supply–demand transactions and is responsible for optimal dispatch of transactions between suppliers and distribution people. It communicates with the retailers to decide the market price and with aggregators to get details of renewable energy power.

2.6 Customer Domain

The main aim of this domain is to permit customers to participate in grid communication. Critical updates from the grid should reach the customer, so communication with customer premises is very vital. The integration of DERs is under the control of this domain. Data exchange taking place between smart meters and energy system interface is also under the supervision of this domain. It interacts with distribution, operations, services, and market domain.

2.7 Service Domain

Delivery of electricity to customers and industries comes under the responsibility of the service domain. It also manages customer billing and customer accounts. Home energy management communication comes under the service domain. It needs to get metering information to process bills. It interacts with operations and the customer domain.

3 Network Types

Different geographical regions are taken care of by constructing an interconnected hierarchical architecture. Individual subnetwork forms part of this network. It is classified into WAN, field area network (FAN), and home area network (HAN). WAN covers the transmission system and large distribution systems. Local SCADA information is captured by intelligent electronic devices (IED) and acts according to the protection commands from control centers. Instruction communication from control centers to electric devices is carried out by WAN. FAN gives a communication facility for small distribution systems to share and exchange information. Home buildings and apartment communications come under the HAN network. Load curtailment, demand response, and load shedding are handled by HAN. It acts as a backbone in customer premise to interface with utility for planning equipment usage and updating solar prices. These network types use power line communication, wireline network, or wireless network as networking technologies. Communication in a substation is shown in Fig. 2.

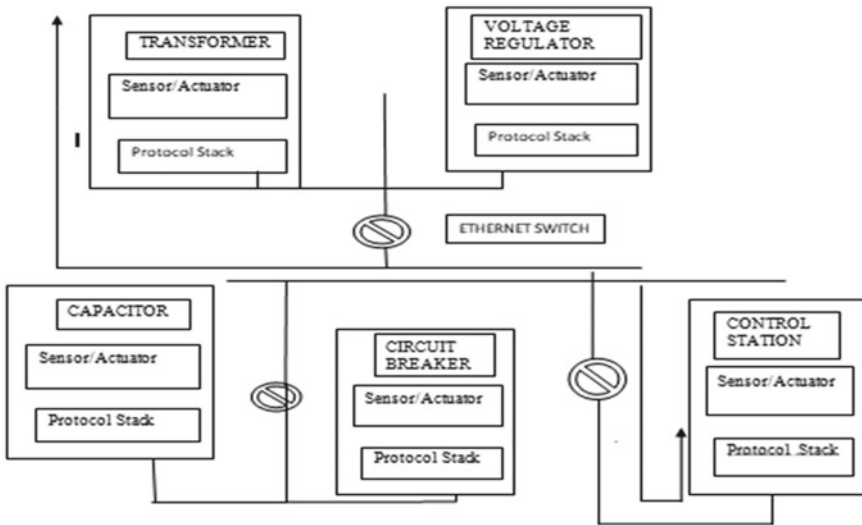


Fig. 2 Communication in a substation

4 Communication Functionality to be Fulfilled

4.1 Wide-Area Situational Awareness (WASA)

Data from electric substation and power lines are collected. Based on this data, wide-area monitoring system, wide-area control system, and wide-area protection system are built. System-level protection is made available by taking the data from the phasor measuring unit. The dynamics due to renewable energy sources are also monitored.

4.2 Distribution Grid Management

Distribution grid needs monitoring of feeders, smart transformers, PEVs, AMU, and demand response (DR) schemes. These applications use FAN within the distribution domain for gathering data.

4.3 Advance Metering Infrastructure

AMI system helps in two-way communication between utility and customers using the Internet of Things technology (IoT). Load forecasting implements the DR scheme. It segregates critical and non-critical loads using FAN and HAN.

4.4 Demand Response

Peak shaving can be implemented using the DR system. Communication protocol enables utility companies to provide dynamic real-time pricing information to the customers to shift the usage of electricity. By providing critical peak pricing and variable critical peak rebates, the customers can be besotted in demand response efforts. End-user interfaces, load control devices, and AMI helps in these communications. HAN is widely used to coordinate the communication of smart appliances.

5 Communication Requirements for Message Transfer

- Network latency
- Data delivery criticality
- Reliability
- Security

- Time synchronization
- Multicast support
- Critical timing requirements.

5.1 Network Latency

Network latency defines the maximum time required for a message to reach its destination. Different messages in a power grid have different network latency. Protection and control messages are critical, and hence, latency requirement has to be met.

5.2 Data Delivery Criticality

High, medium, and non-critical data delivery criticality are used in the smart grid. A high type of delivery scheme is used when an acknowledgment is unavoidable in the end-to-end data delivery. SCADA control commands are of this type. Retrying of message sending is done until a confirmation is obtained. The medium type is used when the receiver can detect data loss and end-to-end confirmation is not required. The source is not concerned whether the data has reached the destination. Measured values of current, voltage, and power are of this type. Critical type is used when data loss is acceptable to the receiver. Messages are sent repeatedly or not.

5.3 Reliability

Reliability stands for resource ampleness and operational security and safety. Electric congestion should be prevented for attaining reliability. Intelligent, self-healing networks facilitate reliable operations of the power grid, increasing grid efficiency.

5.4 Security

The digital infrastructure of the smart grid is susceptible to cyberattacks than the primitive electrical grid. Improving reliability, efficiency, and real-time monitoring of the gadgets can provide cybersecurity. Most modern and flawless security measures must be implemented in the cybernetwork.

5.5 Time Synchronization

Devices in the power grid need to be time-synchronized. In the case of a transmission line running between two substations, if data obtained from substations are not the time-synchronized, comparison is not possible. Devices are installed in substations to provide time synchronization. Protection schemes and WAN require time synchronization.

5.6 Multicast Support

Multiple receivers are there to receive the data. So multicast communication is needed. Person-wise admission control must be provided to reduce data theft.

5.7 Critical Timing Requirements

Communication delay is the time interval between sending and receiving of data. If a trip command is sent to the breaker, it should reach the breaker in 4 msec. Timing varies from country to country.

6 Standards

6.1 IEEE C37.1

This standard provides the basics for the definition, specification, performance analysis, and application of SCADA and automation system. Network requirements on reliability, maintainability, availability, security, expandability, and chargeability are mentioned.

6.2 IEEE 1379

A set of guidelines for communications and interconnection of remote terminal units (RTUs) and intelligent electronic devices (IEDs) in an electric utility substation is bestowed in IEEE 1379 standard. Following these standards eliminate the need for time-consuming and costly interfacing techniques. Implementers need to understand RTU and IED communication standards and the overall concept of the supervisory control and data acquisition (SCADA) system.

6.3 IEEE 1547

IEEE 1547 is a standard for interconnecting distributed resources with electric power systems. It is developed by the Institute of Electrical and Electronics Engineers. It provides prerequisites and specifications needed for performance, operation, testing, safety, and maintenance of the interconnected equipment.

6.4 IEEE 1646

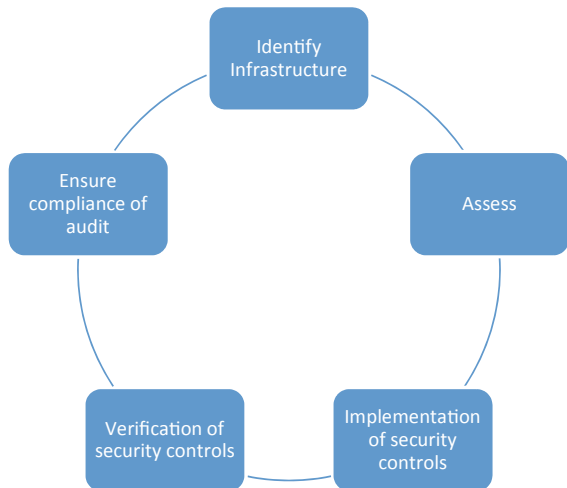
Rules elaborating communication dispatch times of information movement within and external to substation integrated protection, control, and data acquisition systems are described. Communication requirements and system effectiveness in delivering data on time are also mentioned.

7 Cybersecurity

It refers to the protection of networks, hardware, and software from attack, damage, or unauthorized access and rejection of services. Figure 3 shows the duties that have to be carried out in a cybersecurity system loop.

Cybersecurity problems can occur from unknown, competitor, agent of foreign nation, current contractor, former employee, software vendor, hacker, current employee, or malware author. Fraudsters alter electronic data and access credit

Fig. 3 Duties to be carried out in a cybersecurity system



cards or money-related activities. Cybercriminals try to snatch personal information. Hacktivists misuse the computer system for socially and politically motivated gains. Bonnets are Internet-connected devices that are made to infect with malware for data theft, distributed denial of service (DDoS) attack, and credential leakage. Denial of service is the outcome. A network can be penetrated directly or indirectly by using the data obtained from social engineering Web sites, so they need to be banned near the power grids.

The software can be affected due to vulnerabilities present in the communication network. The main attacking strategies are explained below.

7.1 Zero-Day Vulnerabilities

Zero-day vulnerabilities occur when the vendor cannot provide a solution to some weaknesses. At any time, attack can occur. It causes an arms race between hackers and utility. Alarming attacks based on zero-day weaknesses can go unnoticed for a long period. Due to this cyberwarfare, governments are purchasing zero-day exploits from recognized and legalized security companies

7.2 Denial of Service (DoS) and Distributed Denial of Service (DDoS) Attacks

In the advent of the DOS attack, the system cannot react to service requests. It may be the first step for another attack which the hacker plans. TCP SYN flood attack, teardrop attack, smurf attack, ping-of-death attack, and bonnets are examples of DOS attacks.

7.3 Man-in-the-Middle (MitM) Attack

A MitM attack occurs when a hacker inserts itself between the communications of a client and a server. In session hijacking, a session is hijacked. The attacking computer substitutes its IP address for the trusted client, while the server continues the session, believing it is communicating with the client. In IP spoofing of the target, the source receives a packet with the IP source address of a known, trusted host and acts upon it. A replay attack happens when a cyberthief interrupts and saves old messages and then tries to send them later, mimicking one of the participants. The countermeasure is to provide session-wise timestamps or nonce. Encryption and digital certificates are effective safeguarding measures against MitM attacks, thus providing confidentiality and reliability of communications.

7.4 Phishing and Spear-Phishing Attacks

Emails that seem to be from reliable sources are sent to gain control of personal information. It may be in the form of a link to an unauthorized Web site or the form of a malware attack. Spear phishing is a targeted type of activity. Attackers research targets and create messages that are personal and important, and hence, it is difficult to defend. Email spoofing and Web site cloning are typical types of spear phishing. Legitimate Web sites are fooled into entering personally identifiable information (PII) or login credentials. To diminish the extent of being phished, Sandboxing, analyzing email, and critical thinking need to be executed [37–39]. A drive-by download type of attack is a method of spreading malware. Hackers put the seeds of a malicious script into a weak Web site. It takes the loopholes of a Web or operating system. Operating systems need to be updated and never try to visit unknown Web sites that contain un-understandable codes. Try to avoid using unnecessary programs and apps in the device because as the number of plug-ins increases, more vulnerability may creep in.

7.5 Password Attack

Procuring passwords is a frequent and effective type of attack approach. Seizure of password can occur by “sniffing” the connection to the network. Unencrypted passwords are obtained even by wild guessing. Brute-force guessing uses a random approach in which some logic related to the person’s name, family member’s name, job place, or hobbies are applied. In a dictionary attack, common passwords are used to gain control of the user’s computer network. An account lockout policy needs to be implemented to safeguard the system.

7.6 SQL Injection Attack

SQL injection mainly occurs in database-driven Web sites. It happens when a SQL query is executed. SQL commands are lodged into data-plane input. A SQL injector can read and modify sensitive data from the database, and SQL injection operation happens when dynamic SQL is used. The least privileged model of permissions and stored procedures need to be employed in the databases to avoid SQL attack. Input data has to be validated by comparing it with a white list.

7.7 *Cross-Site Scripting (XSS) Attack*

In XSS attacks, the attacker injects venomous JavaScript into the database. The attacker's payload containing page is transmitted to the database which results in the execution of the malicious script. When XSS is employed to exploit additional vulnerabilities, an attacker not only steals cookies, but also log keystrokes, captures screenshots, intrudes into network information, and remotely accesses control of the grid machine. To defend against XSS attacks, developers can clean up data input by users in an HTTP request. Data need to be validated and filtered before echoing. Client-side scripts' scrutinization priority needs to be given to users.

7.8 *Eavesdropping Attack*

Passwords and other personal information can be snatched by eavesdropping attacks. Eavesdropping can be passive or active. Passive eavesdropping information is detected by listening to the message transfer in the network. Active eavesdropping allows a hacker to grab the knowledge by camouflaging as a friendly unit and by giving queries to transmitters. Passive eavesdropping is the starting point of an active attack. Data encryption can prevent eavesdropping.

7.9 *Malware*

Malware is the collective name for several malicious software variants, including viruses, ransomware, and spyware. It is a software type intentionally designed by hackers, for making money illegally or to cause damage to a computer, server, client, or computer network to steal sensitive or confidential information. A computer worm is an autonomous program that replicates itself to spread to other computers and affects multitasking systems connected to a network. It does not require a host or human help for spreading. Trojan horse or trojan is any malware that misleads users of its true intent. It can give the attacker backdoor control over the device and download a virus or a worm, activating a device camera or recording capabilities. This will lead to unwanted popup messages, crashing, and restarting of personal computers.

8 Smart Grid Architecture Model

The model of a smart grid is shown in Fig. 4 [40–42].

To obtain a correlation between devices in the smart grid, interoperability needs to be contemplated on five interoperability layers. The first two layers are related to

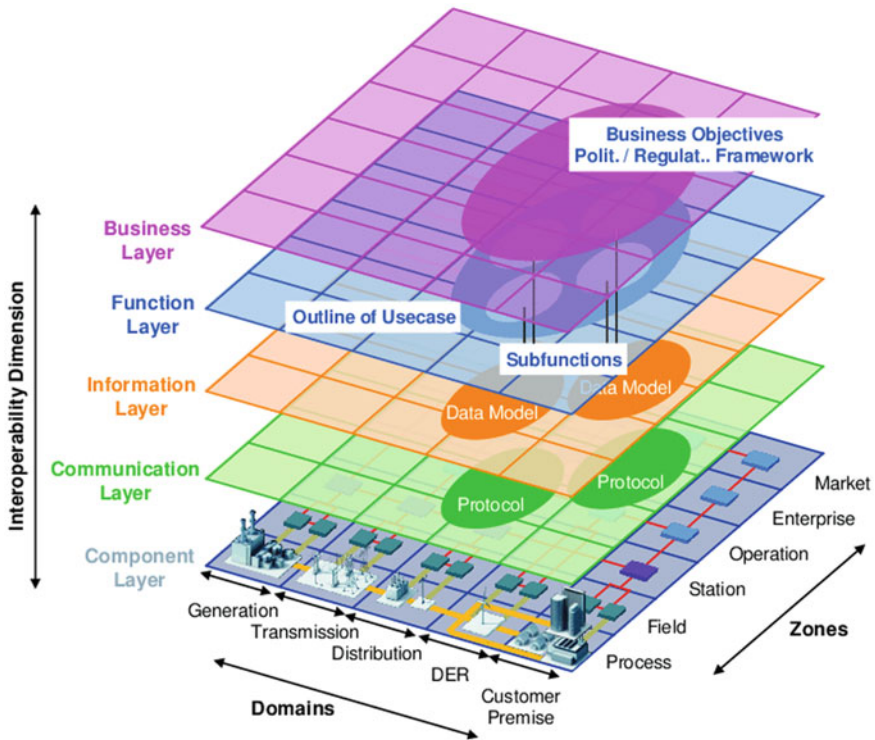


Fig. 4 Smart grid model

functionality, whereas the lower three layers can be associated with the intended technical implementation. Grid-wise architecture council (GWAC) developed the interoperability layers architecture. Business layer provides a business perspective of the information swapping in smart grids. Administration-related and economic infrastructures can be mapped on to this layer. Function layer outlines services and their relationships from an architectural frame of reference. Information layer sketches the fundamental and elementary canonical data models of information among the various components. Component communication layer portrays protocols and procedures for the exchange layer which gives an idea of the physical arrangement of all components including power systems and equipment.

9 Cybersecurity Threats in Smart Grid Advanced Metering Infrastructure

AMI includes heterogeneous devices such as smart meters, collectors, firewalls, routers and gateways, heterogeneous links of power lines, wired, and wireless data

connections. AMI networks have to consider different data stream types such as power usage data control commands, alarms, software patches, and management commands. HAN makes connections to meters. From the meters, a connection is made to the intelligent boxes namely collectors. Routers route the information to the energy provider side. Data delivery in AMI is not policy driven, but it is time driven or request driven, and the AMI network must be accessible from the utility network. Energy usage information and new installation information are collected through HAN. Topological configuration data consist of different links, link properties, logical network zones, and geographical regions. Link definition defines that how devices are connected, the type of algorithms to be used, patches to be installed for acquiring security properties. The encrypted tunnel is provided between the source and destination ports. The service port's priority is provided by administrators.

9.1 Cybersecurity Threats in AMI

9.1.1 Reachability and Link Integrity Threats

Reachability between the meters and receivers should be ensured. Collectors have terminated buffer, and if data tampered, it may turn malicious. Network links need to be monitored and kept intact.

9.1.2 Availability Threats

Improper scheduling results in data loss. Improper scheduling is unknowingly introduced or attackers misconfigure the data resulting in a denial of service. For formal analysis of threats in AMI, the location and values of the errors should be identified accurately, errors need to be modeled and analyzed, and corrective measures have to be implemented. The impact of the threat should be calculated using ease of mitigation, criticality, and other measurable factors. Necessary impact analysis has to be carried out. Modern SAT/SMT solvers and Yices have shown to be powerful and inexpensive both in time and cost to execute millions of constraints for millions of variables in a few seconds.

10 Cybersecurity Solution

All data need to be transferred with confidentiality, integrity, and availability. By applying policies procedures, advanced technologies, education, training, and awareness programs security problems can be reduced.

10.1 Perimeter-Level Protection

This protection is activated within a certain perimeter of WAN. Firewall, intrusion detection system (IDS) which monitors network traffic for illegal activities, harmful activities and breach of policies, intrusion prevention system (IPS) which alarms the network user and removes the harmful component from the network, virtual private network (VPN) that creates an encrypted tunnel between the supplier and end user on top of a public network, antivirus software, role-based access providing schemes, and network admission control (NAC) are the main security providers in this level.

10.2 Interior Security

Firewall, IDS, VPN, AV, host AV, IEEE 1711 (serial connection), network admission control (NAC), and scanning need to be provided within a substation and network to produce defense in depth

10.3 Monitoring

Strict monitoring of the network must be enforced to reduce data leakage and cyberattack.

10.4 Management

Security management measures should be strictly monitored to avoid data hijacking and intrusion. Role-, device-, and location-based access controls reduce security problems.

10.5 Processes

Predefined security rules, barriers, and processes can minimize cybersecurity issues. Security compliance must be made a fundamental necessity in the smart grid.

11 Cybersecurity Standards

11.1 ISO/IEC 27000

General information technology security rules are given in the ISO/IEC 27000 series. International (ISO) and the International Electrotechnical Commission (IEC) jointly published these rules and regulations. It gives a broad perspective of the privacy and clandestinely of IT/technical/cybersecurity issues. Main standards are listed below [37]

- ISO/IE 27000—Information security management systems
- ISO/IEC27004—Information security management monitoring, measurement, analysis, and evaluation
- ISO/IEC27005—Information security risk management
- ISO/IEC 27032—Guideline for cybersecurity
- ISO/IEC 27033—IT network security
- ISO/IE 27038—Specification for digital redaction on digital documents
- ISO/IEC 27039—Intrusion prevention.

Organizations choose to implement these standards for their benefit as they also want to get certified to reassure customers and clients that the rules, recommendations, and regulations have been followed.

11.2 NIST SP 800-82

Security in the industrial control system. The aim of the rules and regulations furnished in the document is to provide directions for securing industrial control systems (ICS), including SCADA systems, distributed control systems (DCS), and other control operations. As per NIST SP 800-82, regulation risk analysis has to be conducted to identify critical parts that need to be segmented to deny or minimize intrusion to sensitive information. Network traffic filtering which uses different technologies to ensure security conditions in various domains and network layer filtering protocols is elaborately explained. State-based filtering which determines the communication hierarchy and application filtering that filters the content of communications between systems at the application layer is also given special consideration.

11.3 IEC 62351

IEC 62351 is a standard developed by WG15 of IEC TC57. This is developed for handling the security of TC 57 series of protocols including IEC 60870-5 series, IEC 60870-6 series, IEC 1850 series, IEC 61970 series, and IEC 61968 series. The different security objectives include authentication of data transfer through digital signatures, ensuring only authenticated access, prevention of eavesdropping, prevention of playback and spoofing, and intrusion detection.

The main aim of IEC 62351-10—Security Architecture standards are explained below

- Description of security structure for the IT infrastructure
- Finding the location of crucial points such as substation control center and substation automation centers
- Identification of suitable security measures
- Applicability of IT domain standards.

11.4 The North American Electric Reliability Corporation Critical Infrastructure Protection (NERC CIP)

NERC developed CIP standards to exhort utilities to develop an incipient set of security measures. It contains 9 standards and 45 requirements which cover the security of electronic gadgets and the safeguarding of critical cyber assets and as personnel and training, security management, and disaster recovery planning. NERC CIP recommends the utilization of firewalls to block malicious ports and the employment of cyberattack monitoring tools. Organizations need to enforce IT controls protecting access to critical cyberassets. Penalties for recalcitrance with NERC CIP include fines and license cutting.

11.5 NISTIR 7628 (Security in Power System)

Smart grid use data for informed decision making. The major task involves state estimation, stability assessment, energy forecasting, and situational awareness. NISTIR 7628 provides a master plan and guidelines for formulating cybersecurity strategies for different grid-related organizations. A detailed description of various smart grid domains, 22 logical interface categories, and their security specifications is vividly presented. Secret key, public key, and hash function type cryptographic details are elaborated with potential alternatives. Rules and regulations are incorporated to study privacy impact assessment, and the problems that creep into the system as new capabilities are incorporated. Special mention is given for R&D themes that identify where the state of the art falls short of meeting the envisaged functional, integrity,

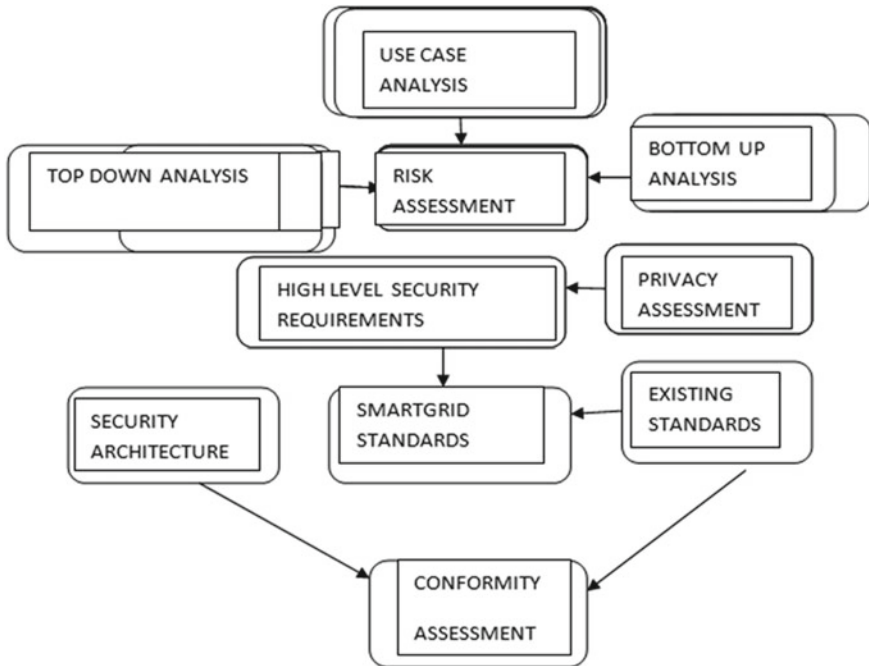


Fig. 5 Tasks in a cybersecurity system

reliability, scalability, and security requirements of the smart grid. Various tasks in the cybersecurity strategy of the power grid are shown in Fig. 5 [36].

The selection of use cases provides an envelope for developing risk assessment and logical models. Risk assessment recognizes culpability, threats, and assets. Collaborative work of information technology personals and power system personals help to develop security requirements in which top-down analysis and bottom-up analysis are carried out. Logical communications interfaces are identified by the logical reference model. Loopholes in communication standards are identified, and corrective measures are developed. Testing and certification come under the conformity assessment.

Security requirements section contains the proposed security requirements for the smart grid [36]. The recommended security requirements are categorized into families primarily based on NIST SP 800-53. The following information is included with each security requirement: (1) Security requirement identifier and name. Three main components in the identifier are (a) SG—for smart grid, (b) the family name, e.g., AC for access control and CP for continuity of operations. (c) unrepeated numeric identifier, SG.AC-3 and SG.CP-2. (2) Category which checks whether the security requirement is a governance risk compliant (GRC), common technical or unique technical requirement. For common technical security requirements, confidentiality,

integrity, and availability standards are listed. (3) The requirement defines security-related measures to be carried out by the smart grid information system. (4) Additional information necessary for understanding is published in supplemental guidance section (5) The requirement enhancements portray the security capability to (i) build additional functionality in a requirement and/or (ii) escalate the strength of a requirement. Requirement intensifications are numbered sequentially in each requirement. (6) Additional considerations provide extra statements of security potentiality that can be used to increase the security requirement. (7) The impact-level allocation focuses on the security requirement and requirement boosting, at each impact level.

12 Current Research Works and Future Scope

Current research works include simulation and validation of self-healing resilient smart power grid, game theory-based cyberattack modeling, and false data injection attack. Game theory is a branch of mathematics that provides numerous mathematical tools for strategic decision making in the cybersecurity field. Gambit is an open-source tool for building, analyzing, and exploring game models. The basic elements of a game are players, strategies, and payoffs. A game is an elaboration of a strategic circumstance in which two or more players are included. Each player has preferences among all the possible results. A brilliant game designer provides continuous challenges to his players, each of which ends in another challenge. Presently game theory is used in power systems for modeling cyberattacks and energy routing in a smart grid network.

A system that intelligently uses knowledge, sensing, and control, and communication know-how for real-time monitoring, rapid isolation, and immediate restoration of supply in a power grid is known as a resilient, self-healing grid. Phasor measurement units (PMU) which provide accurate GPS-based time stamping polls data from various grid devices, so that instantaneous analysis is possible. This helps to prevent the spread of disruption and reduce the number of outages. It is therefore critical to protect PMU networks against cyberthreats and system faults. The emerging software-defined networking (SDN) technology can be used to model and configure resilient network and self-healing algorithms against cyberattacks. In the event of a cyberattack, the PMU network is made to reconfigure itself to reroute measurement data for preserving the power system observability.

13 Conclusions

Cybersecurity issues are alarmingly increasing due to the rapid growth of cyber technology. The latest technological inventions can be used to circumvent cybersecurity issues.

Acknowledgements I would like to keep on record the guidance and suggestions given by Dr. Shantishwarup, Professor, Department of Electrical Engineering, IIT Madras, Chennai, and Dr. Balakrishna Pamulaparthy, Senior Lead R&D Engineer, GE Renewable Energy, Hyderabad Technology Centre.

References

1. J. E. Dagle, Cyber security of the electric power grid, in *Proceedings ISGT* vol. 1, (2012) pp. 1–2
2. P. Kaster, P. K. Sen, Power grid cyber security: challenges and impacts, in *2014 North American Power Symposium*
3. J. Jarmakiewicz, K. Parobczak, K. sSlanka, Cybersecurity protection for power grid control infrastructures. *Int. J. Critic. Infrastruct. Protec.*
4. D.B. Rawat, C. Bajracharya, Cybersecurity for smart grid systems: Status, challenges, and perspectives, in *SoutheastCon* (2015), pp. 1–6
5. S. Shapsough, F. Qatan, R. Aburukba, F. Aloul, A. Al Ali, Smart grid cybersecurity: challenges and solutions, in *International Conference on Smart Grid and Clean Energy Technologies (ICSGCE)* (2015), pp. 170–175
6. X. Liang, K. Gao, X. Zheng, T. Zhao, A Study on Cyber Security of Smart Grid on Public Networks, in *IEEE Green Technologies Conference* (2013) pp. 301–308
7. M. Essaaidi, et al., An overview of smart grid cyber-security state of the art study, in *3rd International Renewable and Sustainable Energy Conference (IRSEC)* (2015) pp. 1–7
8. W. Wang, Z. Lu, Cybersecurity in the smart grid: survey and challenges. *Comput. Netw.* **57**(5), 1344–1371 (2013)
9. G. Leon, Smart planning for smart grid AMI mesh networks. Technology white paper. EDX Wireless [Online]. Available: http://www.edxwireless.biz/news/EDXWP_Smart_Grid_AMI_Mesh_Networks_May_11.pdf
10. The U.S. Pacific Northwest National Laboratory (PNNL). AMI communication requirements to implement demand-response: applicability of hybrid spread spectrum wireless [Online]. Available: https://www.pnnl.gov/main/publications/external/technical_reports/PNNL-20806.pdf
11. Y. Jiang, C.C. Liu, M. Diedesch, E. Lee, A.K. Srivastava, Outage management of distribution systems incorporating information from smart meters. *IEEE Trans. Power Syst.* **31**(5), 4144–4154 (2016)
12. A. Hari, S. Suri, G. Parulkar, Detecting and Resolving Packet Filter Conflicts
13. H. Hamed, E. Al-Shaer, W. Marrero, Modeling and verification of IPSec and VPN security policies, in *13th IEEE International Conference Network Protocols (ICNP'05)* (2005). p. 10
14. E.S. Al-Shaer, H.H. Hamed, Discovery of policy anomalies in distributed firewalls. *IEEE INFOCOM* **4**(2004), 2605–2616 (2004)
15. L. Yuan, H. Chen, J. Mai, C.N. Chuah, Z. Su, P. Mohapatra, Fireman: a toolkit for firewall modeling and analysis, in *2006 IEEE Symposium Security and Privacy (S&P'06)*, (CA: Berkeley/Oakland, 2006)
16. R.C.W. Phan, Authenticated modbus protocol for critical infrastructure protection. *IEEE Trans. Power Del.* **27**(3), 1687–1689 (2012)
17. G. Hayes, K. El-Khatib, Securing modbus transactions using hash-based message authentication codes and stream transmission control protocol, in *2013 Third International Conference on Communications and Information Technology (ICCIT)*, (Beirut, 2013), pp. 179–184

18. Modbus Application Protocol Specification, V1.1B Modbus Organization; 2006. Available: <https://www.modbus-IDA.org>
19. C.W. Ten, C.C. Liu, G. Manimaran, Vulnerability assessment of cybersecurity for SCADA systems. *IEEE Trans. Power Syst.* **23**(4), 1836–1846 (2008)
20. N. Falliere, L.O. Murchu, E. Chien, W32.Stuxnet Dossier. Symantec security response, Version 1.4; Feb. 2011. Available: https://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf
21. Y. Zhang, L. Wang, Y. Xiang, C.W. Ten, Inclusion of SCADA cyber vulnerability in power system reliability assessment considering optimal resource allocation. *IEEE Trans. Power Syst.* **31**(6), 4379–4394 (2016)
22. V. Namboodiri, V. Aravinthan, S.N. Mohapatra, B. Karimi, W. Jewell, Toward a secure wireless-based home area network for metering in smart grids. *IEEE Syst J* **8**(2), 509–520 (2014)
23. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive Data Analysis for Energy Management of a Smart Factory Leading to Sustainability, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, ed by Favorskaya, M., Mekhilef, S., Pandey, R., Singh, N. vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
24. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
25. R.N. Shaw, P. Walde, A. Ghosh, A new model to enhance the power and performances of 4 × 4 PV arrays with puzzle shade dispersion. *Int. J. Innovat. Technol. Explor. Eng.* **8**(12) (2019)
26. Y. Belkhier, A. Achour, N. Ullah, R.N. Shaw, Modified passivity-based current controller design of permanent magnet synchronous generator for wind conversion system. *Int. J. Model. Simul.* (2020). <https://doi.org/10.1080/02286203.2020.1858226>
27. X. Liang, X. Li, R. Lu, X. Lin, X. Shen, UDP: usage-based dynamic pricing with privacy preservation for the smart grid. *IEEE Trans. Smart Grid* **4**(1), 141–150 (2013)
28. B. Krebs FBI: Smart meter hacks likely to spread; 2012. Available: HTTP://
29. R. H. Danville, Utilities see an increase in meter tampering; 2012. Available: <https://www.wset.com/story/20442252/danville-utilities-sees-increase-in-meter-tampering>
30. North American Electric Reliability Corporation (NERC). CIP standard. Available online: ftp://www.nerc.com/pub/sys/all_updl/standards/sar/CIP-002-009-1_30-day_Pre-ballot_Comment.pdf
31. National Institute for Standards and Technology. Guidelines for smart grid cyber security, NISTIR 7628. [Online]. Available: <https://online.wsj.com/news/articles/SB10001424052702304851104579359141941621778>.
32. National Institute for Standards and Technology. The cybersecurity coordination task group: smart grid cybersecurity strategy and requirements. Available online: https://csrc.nist.gov/publications/nistir/ir7628/nistir-7628_vol2.pdf
33. O. Khaled, A. Marín, F. Almenares, Analysis of secure TCP/IP profile in 61850 based substation automation system for smart grids
34. A.A. Elbaset, Y.S. Mohamed, A.N.A. Elghaffar, IEC 61850 communication protocol with the protection and control numerical relays for optimum substation automation system. *J. Eng. Sci. Technol. Rev.* **13**(2), 1–12. 12 (2020)
35. H. León C. Montez, O. Valle, F. Vasques, Real-time of time-critical messages in IEC 61850 electrical substation
36. Shailesh Chauhan
37. D.P. de Mattos, et al., IEC 61850 packet generator for testing substation communication”
38. M. Mekkanen, K. Kauhaniemi, Wireless light-weight IEC 61850 based loss of mains protection for smart grid
39. V.S. Rajkumar, M. Tealane, A.S. Tefanov, P. Palensky, Cyber attacks on protective relays in digital substations and impact analysis

40. L. Sathyadevan, J. Starck, Digitized copper—advantages, disadvantages, and outlook of a power and process integrated control system using IEC 61850
41. [nvlpubs.nist.gov > nistpubs > ir > NIST.IR.7628.pdf](http://nvlpubs.nist.gov/nistpubs/ir/NIST.IR.7628.pdf)
42. J. Trefke, S. Rohjans, S. Lehnhoff, L. Nordström, A. Saleem Smart grid architecture model use case management in a large European smart grid project. Innovative Smart Grid Technologies Europe (ISGT EUROPE), 2013 4th IEEE/PES

Innovative Approach to Harvest Energy from Traffic-Induced Deformation



Nilimamayee Samal, O. J. Shiney, and Abhinandan Patra

Abstract This article targets on the development of an efficient green energy source to power the wireless sensor nodes and small power electronic devices used in the road ways and traffic system. A simulation-based model of Piezo-electric Traffic Energy Harvesting (PTEH) System has been presented to advocate the idea of scavenging electric energy from vehicular vibrations. The viability of the system is investigated using the real time MATLAB and Simulink platform. The material property of the piezo-electric sensors and the electronic factors associated with the additional components used to achieve the maximum kinetic-energy utilization were taken into consideration for proposing the structural design of the system. In addition to the designing, this paper estimates the total energy generation using MATLAB and Simulink platform. The environment-friendliness of the proposed system has been investigated by doing the carbon emission analysis.

Keywords Piezo-stack sensor · Energy harvesting · Traffic induced deformation · Vibration

N. Samal (✉) · O. J. Shiney
School of Electrical, Electronics and Communication Engineering, Galgotias Universty, Delhi,
India

A. Patra
IBM, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_4

1 Introduction

The global demand for usable power is jumping up day-by-day. The developed countries are carrying on with consuming more and more energy, while developing countries' demand is rising gradually. As per the "International Energy Agency's 2019 World Energy Outlook", if the developed and developing countries continue to move in the present track without changing the policy, the demand of usable energy will rise by 1.3% per year till 2040.

The solution for fulfilling the rising energy demand lies in harvesting more and more energy from the unutilized energy sources. The classification of unutilized energy sources is shown in Fig. 1. The scientists are focusing on the area of energy harvesting from natural energy sources for more than a decade. Many models/devices have been developed for clean energy harvesting from different environmental sources, vehicular sources, industrial sources, human motions etc. These energy harvesting systems have been demonstrated using magneto-electric, thermoelectric, piezoelectric transducers etc.

Piezoelectric material generates the electrical energy when it is under mechanical stress, vibration, force, pressure etc. The energy conversion property of piezoelectric

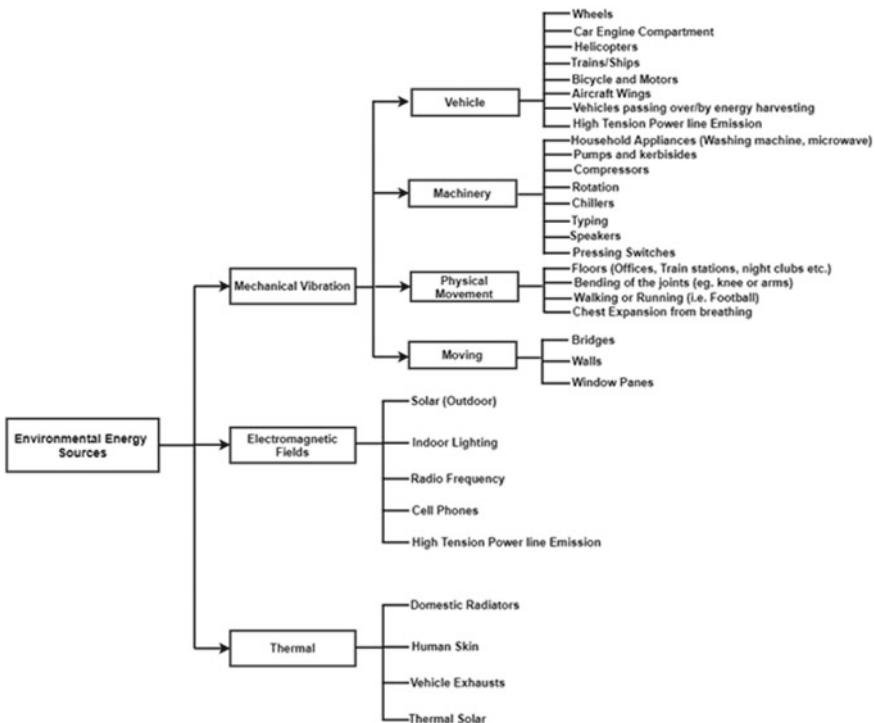


Fig. 1 Classification of environmental energy source

transducers makes it suitable for energy harvesting applications. Different types of piezo-electric materials used in energy harvesting process are; single-crystal, lead-based piezo-ceramics, lead-free piezo-ceramics, piezo-polymers etc. Depending on the configuration, the piezoelectric transducers are classified as:

- Cantilever beam type
- Diaphragm type
- Cymbal type
- Stack type.

In this manuscript, Piezo-stack sensor has been used in the model for conversion of vehicular deformation into electrical energy.

2 Literature Review

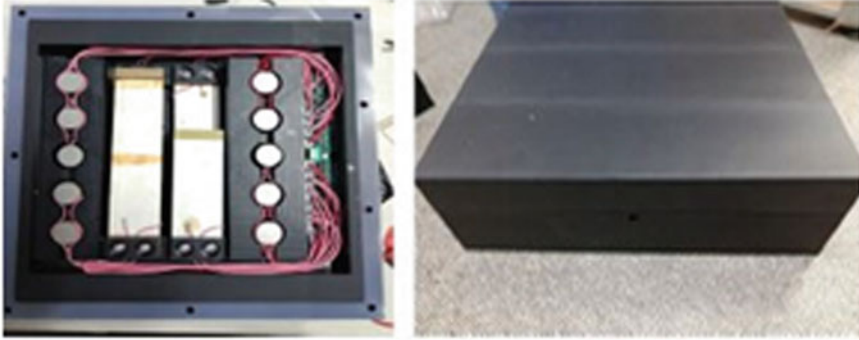
The vehicular vibration energy eventually destroys the pavement structures and it is not easy to collect that energy. The mechanical to electrical conversion property of piezoelectric transducer are used to convert the vehicular vibration energy into electrical energy. This electrical energy can be used to feed power to the road side electric appliances, such as traffic signal lights, advertising boards etc.

In 2010, Prof. H. Aramovich (CEO of Innowattech) and Associate Prof. of Technion Institute of Technology headed a project on piezoelectric energy harvesting (PEH) in roads of Israel. They observed that when the piezoelectric energy generators are installed 6 c.m. beneath the road level maintaining 30 c.m. distances from each other, the system is able to produce power of 400 kW in one hour for a 1-km stretch assuming the traffic of 600 vehicles/h [1]. Now-a-days PEH systems have been implanted in many office corridors, mostly in Japan, to lighten up LEDs when it senses the footsteps.

Najini and Muthukumarswamy presented a simulation-based model to advocate the idea of harvesting energy from road traffic using piezo-electric material. From this simulation work, they observed that energy of 137, 255 and 469 kWh can be yielded from a single lane road with vehicle speed of 80,100 and 120 km/h respectively assuming the traffic of 500 vehicles per hour [2]. Jasim et al. have also reported numerical simulation model of PEH system for roadways applications [3]. Xu et al. conducted an experimental study on piezoelectric roadways energy harvesting systems. They concluded that out of PZT-4, PZT-8 and PZT-5H (all at the same size), the PZT-5H possess a high voltage, high relative dielectric constant and high capacitance when load frequency is more than 5 Hz. It is able to produce comparatively more power under same load. Hence, it is more suitable to be used in piezo-electric boxes for pavement energy harvesting [4]. The dielectric properties of PZT-4/8/5H are listed in Table 1 (Data taken from [4–7]). Yang et al. designed a piezo-electric power generation unit using stacked array piezo-electric transducer with MC nylon as packaging material for road vibration energy harvesting. They proved the practical significance of the system by performing on-site test (Fig. 2)

Table 1 Properties of PZT-4H, 8H and 5H [5]

PZT material	ϵ_{33}	d_{33} (10^{-12} coulomb/newton)	K	Curie's Temp. ($^{\circ}\text{C}$)
PZT-4	1300.00	250.00	0.54	300.00
PZT-8	1020.00	220.00	0.50	310.00
PZT-5H	2000.00	410.00	0.60	260.00

**Fig. 2** Piezoelectric pavement energy harvesting box [6]

[8, 9]. Rodriguez et al. used PZT and lead-free PIC-700 for designing traffic-energy-harvesting system. They concluded that the relative error in practical implementation of the theoretical model is around 3% in both the ways, by using PZT and PIC-700 [10].

3 Methodology

Three different models are involved in the simulation of proposed PTEHS model, viz. Piezo-stack Sensor model, Vehicle Model and Tire and Road Dynamics model. The Piezo-stack model realizes the optimum conditions/parameters for maximum energy output of PTEHS. The optimized conditions are obtained by gradual (step-by-step) refining of this model. This gradual refinement was done by varying the frequencies in three different phases, viz. primary, extended and optimum frequency in first, second and third cycle respectively. The gradual refinement of the piezo-stack model is shown in Fig. 3.

In the proposed model P-5E piezo-stack has been selected, as it has very promising piezo-electric characteristics and its energy conversion efficiency is very high. The characteristics/properties of the selected piezo-stack [P-5E, material- Lead Titanate, Lead Zirconate ($\text{PbTiO}_3 \cdot \text{PbZrO}_3$)] has been depicted in Table 2.

The parameters associated with the piezo-stack used in the simulation work are depicted in Table 3.

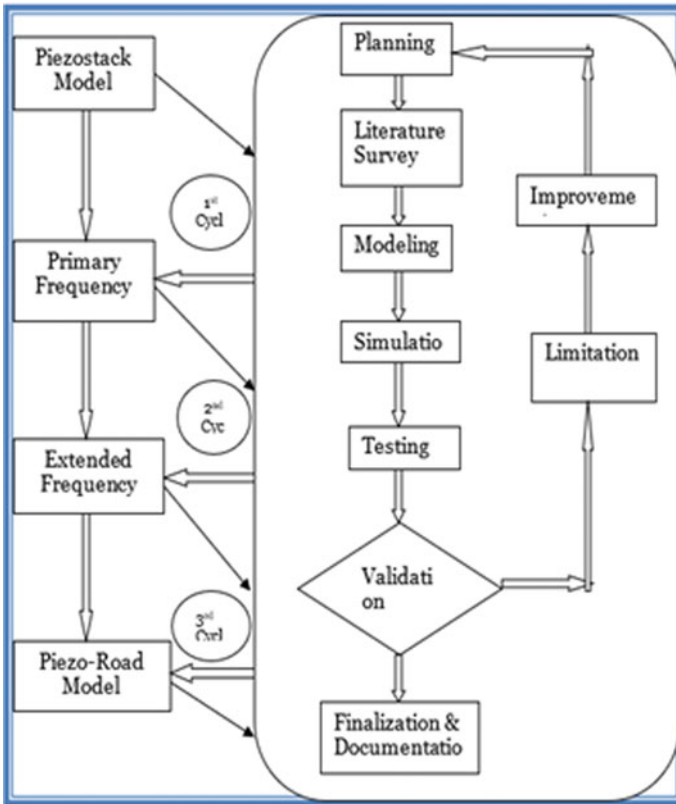


Fig. 3 Gradual refinement of piezo-stack model

Table 2 Properties of P-5E

Property	Symbol	Value
Piezo-electric constant	d_{33}	271×10^{-12} m/V
Electro-mechanical coupling factor	k_{33}	62%
Relative dielectric constant	$\epsilon_{33}^T / \epsilon_0$	1510
Mechanical Q	Q_M	970

4 Design Implementation of Simulink Model

- The PTEHS Model was executed using MATLAB-Simulink Platform. Three different models are involved in the simulation, viz. Piezo-stack Sensor model, Vehicle Model and Tire and Road Dynamics model. The modelling steps of PTEHS are represented in Figs. 4a–d), 5 and 6. The deformation on the pathway due to vehicular load has been calculated using Eq. (1), which is obtained from

Table 3 Parameters associated with piezo-stack used

Parameter	Value
Stack area	100 mm ²
Stack length	36 mm
No load displacement at V ₀ V	0.038 mm
Piezo layer thickness	0.3 mm
Test voltage (V ₀)	3.8e+03 N
Blocking force at V ₀ V	120 V
No. of layers	50
Piezoelectric charge constant	5e−10 m/V
Dielectric constant	2.124e−08 F/m
Elastic compliance	1.9e−11 m ² /N
Capacitance	13 uF
Damping	2 N(m/s)
Effective mass	50 g

Love Plate Theory (2). Tire Dynamics, Vehicle weight and inertia, Vehicle weight and inertia, Load Transfer cases and Aerodynamics are considered for vehicle model. The Tire Dynamics is executed Longitudinal Tire model. The Vehicle Load model is represented in 3 Dimensional degrees of freedom with suspension and damper characteristics.

$$D\nabla^4\omega(x, y, t) + \rho h \frac{\partial^2\omega(x, y, t)}{\partial t^2} + k\omega(x, y, t) = F(x, y, t) \quad (1)$$

$$\nabla^4\omega(x, y, t) + \frac{\partial^4\omega(x, y, t)}{\partial x^4} + 2\frac{\partial^4\omega(x, y, t)}{\partial x^2\partial y^2} + \frac{\partial^4\omega(x, y, t)}{\partial y^4} \quad (2)$$

5 Results and Discussion

For estimation of total power generation by deducing the roadway deformation dynamic response, an equation based mathematical model is represented by (3) The co-efficient of friction between the road and the tire has been taken into consideration for calculating the road deformation, which causes pressure on the piezo-stack sensor. The displacement/deformation and the piezo-electric property have been used to determine the output voltage and energy generated per piezo-stack sensor by incorporating Eq. (4).

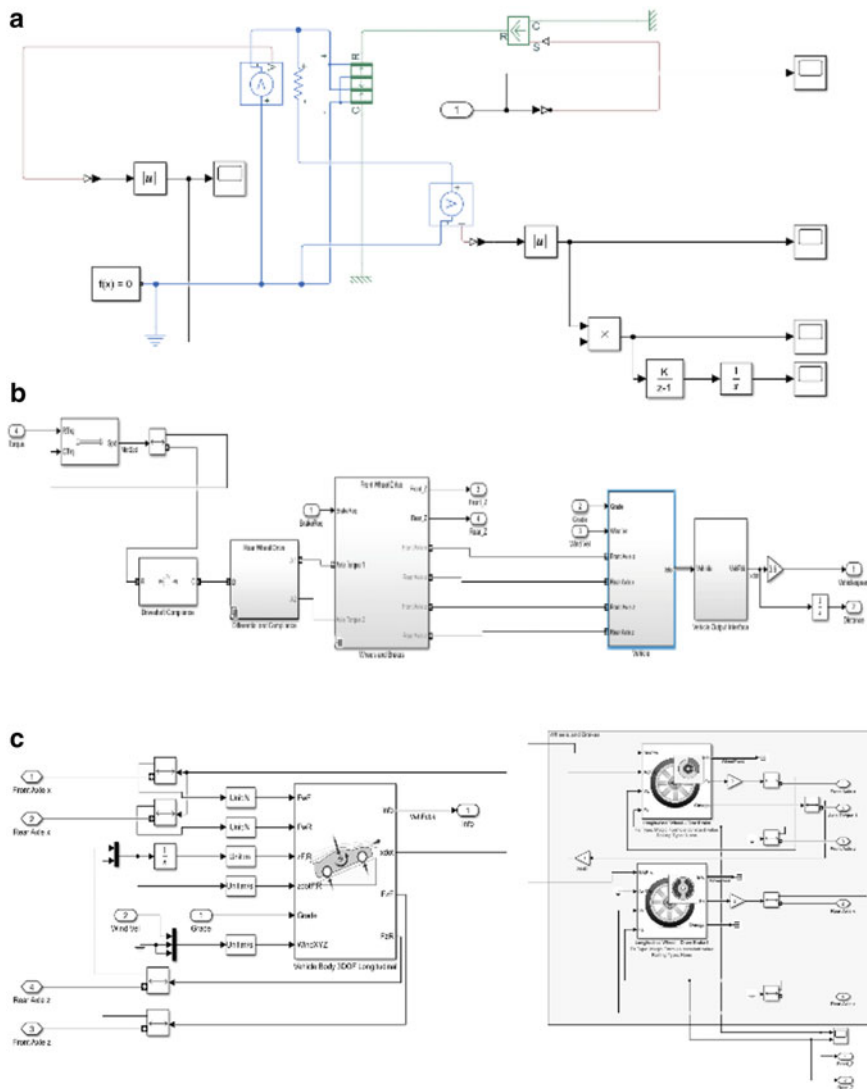


Fig. 4 a Piezoelectric sensor model, b Vehicle model, c Tire and road Dynamics model (Part-1), d Tire and road dynamics model (Part-2)

$$\Pi \frac{\partial^4 \omega_m(x,t)}{\partial x^4} - 2 \frac{m\pi}{b} \frac{\partial^2 \omega_m(x,t)}{\partial x^2} + \left(\frac{m\pi}{b}\right)^4 \omega_m(x,t) + \rho h \frac{\partial^2 \omega_m(x,t)}{\partial t^2} + k \omega_m(x,t) = f_m(x,t) \quad (3)$$

$$C_o \frac{dV(t)}{dt} + \frac{V(t)}{R} = \frac{dQ(t)}{dt} \quad (4)$$

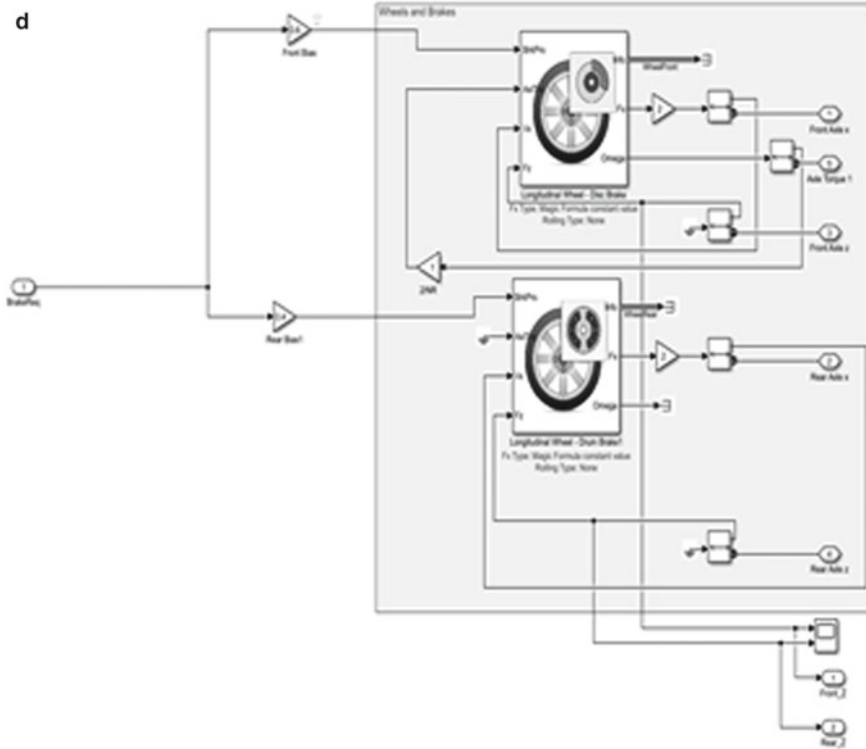


Fig. 4 (continued)

The PTEH System is not involved in any Green House Gas Emission directly. But the system is totally dependent on the vehicles which are fed petrol, diesel or any source of fuel. Tables 4 and 5 shows the estimation of CO₂ in the proposed system. The following assumptions are considered for the estimation of Carbon Di-oxide emission per day.

- The traffic rate varies from 25 to 500 vehicles per hour.
- Vehicles present are only Maruti Suzuki Swift and Toyota Kirloskar Motor (For estimation of average CO₂ emission, the vehicles which emit maximum and minimum CO₂ has been considered)

The estimated carbon emission per annum is only 165 tonnes for the proposed system, which is under permissible limit.

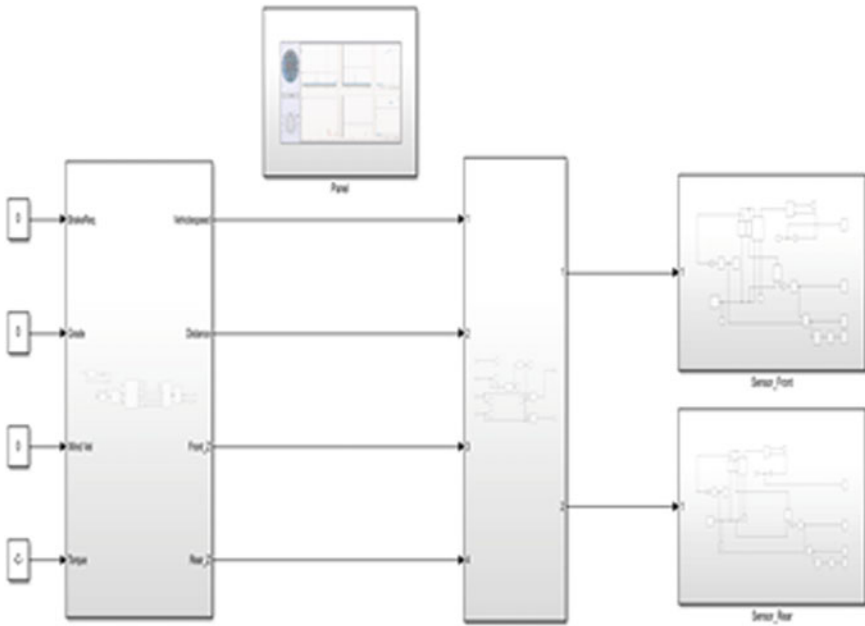


Fig. 5 Overall simulink model for generation of electrical energy from deformation caused by traffic

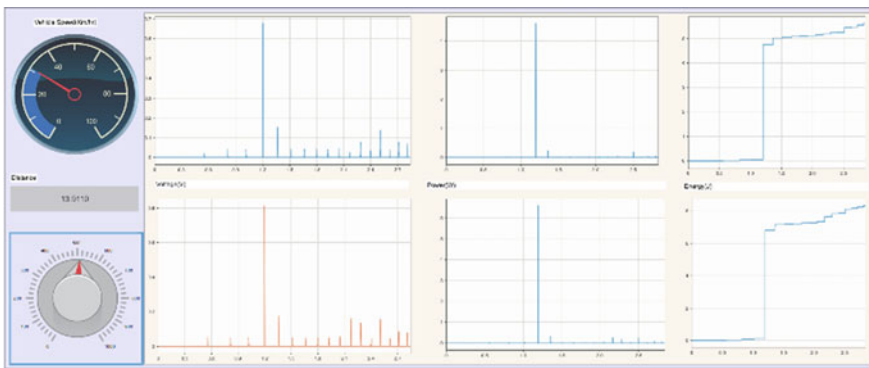


Fig. 6 Voltage, power and energy output shown on control panel

6 Conclusions

This article emphasizes on generation of sustainable energy from the vehicular sources. From the above model, the energy generated due to passing of a single vehicle on a single piezo-stack sensor at the speed of 30 km/h was found to be 12 J (5 and 7 J due to exertion of pressure by the front and rear wheel respectively). The

Table 4 Estimation of carbon emission

Traffic	Traffic rate/h	Duration (h)	Name of the vehicle	Rate of CO ₂ Emission (g/km) per vehicle	CO ₂ Emission (kg/h)	CO ₂ Emission (kg/day)
Rash hour	500	6	Maruti Suzuki Swift	109.5	67.55	
			Toyota Kirloskar Motor	160.7		
						648.48
			Maruti Suzuki Swift	109.5		
Off hour	100	18	Toyota Kirloskar Motor	160.7	13.51	
Rash hour	500	6	Maruti Suzuki Swift	109.5	67.55	
			Toyota Kirloskar Motor	160.7		
						891.66
			Maruti Suzuki Swift	109.5		
Off hour	200	18	Toyota Kirloskar Motor	160.7	27.02	
Rash hour	200	6	Maruti Suzuki Swift	109.5	27.02	
			Toyota Kirloskar Motor	160.7		
						222.915
			Maruti Suzuki Swift	109.5		
Off hour	25	18	Toyota Kirloskar Motor	160.7	3.3775	

(continued)

Table 4 (continued)

Traffic	Traffic rate/h	Duration (h)	Name of the vehicle	Rate of CO ₂ Emission (g/km) per vehicle	CO ₂ Emission (kg/h)	CO ₂ Emission (kg/day)
Rash hour	300	6	Maruti Suzuki Swift	109.5	40.53	
			Toyota Kirloskar Motor	160.7		
						364.77
			Maruti Suzuki Swift	109.5		
Off hour	50	18	Toyota Kirloskar Motor	160.7	6.755	
Rash hour	200	6	Maruti Suzuki Swift	109.5	27.02	
			Toyota Kirloskar Motor	160.7		
						283.71
			Maruti Suzuki Swift	109.5		
Off hour	50	18	Toyota Kirloskar Motor	160.7	6.755	
Rash hour	400	6	Maruti Suzuki Swift	109.5	54.04	
			Toyota Kirloskar Motor	160.7		
						567.42
			Maruti Suzuki Swift	109.5		
Off hour	100	18	Toyota Kirloskar Motor	160.7	13.51	

(continued)

Table 4 (continued)

Traffic	Traffic rate/h	Duration (h)	Name of the vehicle	Rate of CO ₂ Emission (g/km) per vehicle	CO ₂ Emission (kg/h)	CO ₂ Emission (kg/day)
Rash hour	300	6	Maruti Suzuki Swift	109.5	40.53	
			Toyota Kirloskar Motor	160.7		
Off hour	100	18	Maruti Suzuki Swift	109.5	13.51	
			Toyota Kirloskar Motor	160.7		

energy generated is dependent on the no. of piezo-stack sensor embedded beneath the road, the no. of vehicles passing on it, the speed of the vehicle and the load intensity of the vehicle. In a developing country like India, where the traffic rate is very high, the proposed model will be very beneficial for harvesting useful electrical energy from the unused vehicular deformations.

The carbon emission rate of the proposed system is under permissible limit. The developments in the smart technology and smart vehicles may nullify the carbon emission in the proposed system in future, as the smart vehicles use electricity to travel.

Table 5 Estimation of Annual Carbon Emission

Months	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Rush traffic rite	500	300	400	300	300	200	200	200	300	500	500	500
Off-Traffic Rate	100	100	100	50	50	50	25	50	50	100	200	200
No. of days	31	28	31	30	31	30	31	31	30	31	30	31
CO ₂ Emission/day (kg/day)	648.48	486.36	567.42	364.77	364.77	283.71	222.915	283.71	364.77	648.43	891.66	891.66
CO ₂ Emission/Month (kg/day)	20,102.88	13,618.03	17,590.02	10,943.1	11,307.87	83,511.3	6910.365	8795.01	10,943.1	20,102.88	26,749.8	27,641.46
CO ₂ Emission/Annun (kg/day)	165.123865 ≈ 165											

References

1. Z. Zhang, H. Xiang, Z. Shi, Modeling on piezoelectric energy harvesting from pavements under traffic loads. *J. Intell. Mater. Syst. Struct.* SAGE Publ. **27**(4), 567–578 (2015). <https://doi.org/10.1177/1045389X15575081>
2. H. Najini, A. Muthukumaraswamy, Piezoelectric energy generation from vehicle traffic with technoeconomic analysis. *J. Renew. Energy*, Hindawi **2017**, 1–16 (2017). <https://doi.org/10.1155/2017/9643858>
3. A. Jasim, G. Yesner, H. Wang, A. Safari, A. Maher, B. Basily, Laboratory testing and numerical simulation of Piezoelectric energy harvester for roadway applications. *Appl. Energy*, Elsev. **224**, 438–447 (2018). <https://doi.org/10.1016/j.apenergy.2018.05.040>
4. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
5. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol. 661. ed by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
6. R.N. Shaw, P. Walde, A. Ghosh, A new model to enhance the power and performances of 4x4 PV arrays with puzzle shade dispersion. *Int. J. Innovat. Technol. Explor. Eng.* **8**(12) (2019)
7. X. Xu, D. Cao, H. Yang, M. He, Application of piezoelectric transducer in energy harvesting in pavement. *Int. J. Pavement Res. Technol.* **11**(4), 388–395 (2017). <https://doi.org/10.1016/j.ijprt.2017.09.011>
8. H. Yang, L. Wang, Y. Hou, M. Gou, Z. Ye, X. Tong, D. Wang, Development in stacked-array-type piezoelectric energy harvester in asphalt pavement. *J. Mater. Civ. Eng. ASCE Library* **29**(11), 1–9 (2017). [https://doi.org/10.1061/\(ASCE\)MT.1943-5533.0002079](https://doi.org/10.1061/(ASCE)MT.1943-5533.0002079)
9. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON) Held at Deenbandhu Chhotu Ram University of Science and Technology, Sonapat, India on FEB 28 - March 1 2020*
10. M.V. Rodriguez, F.J. Jimenez, L. Pardo, P. Choa, A.M. Gonzalez, J. de Frutos, A new prospect in road traffic energy harvesting using lead-free piezoceramics. *Mater. PMC* **12**(22), 1–16 (2019). <https://doi.org/10.3390/ma12223725>

Impact Analysis of Single Line to Ground Fault on Industrial Loads Using Typhoon HIL



R. Rajesh Kanna, R. Raja Singh , and D. Arun Dominic

Abstract The demand for electrical power is now increasingly growing day by day, which results in the transition of enormous power by increasing the ability of the transmission line from one location to another. As a result of the development of the structure of the power grid, detecting power system transmission line faults every year becomes a challenging task as the system becomes more complicated. Controlling and monitoring of power systems are inevitable due to the constant growth of the power system network. In order to find the correct decision, to search for any possible problems and fault conditions in a broad interconnected network, modeling and analysis of the power system have become significant. On the overhead transmission line, a fault will rise while the two or more of transmission lines touch each other. Generally, the faults in overhead transmission lines can be categorized as the line to ground fault, the line to line fault, double line to ground fault and three-phase faults. Single line ground faults mostly occur in the distribution network. The main purpose of this work is to study or explore the single line to ground fault and recognize the impact of the transmission line fault along with the industrial load under different loading conditions that is joined with transmission line according to the distance variation using Typhoon HIL.

Keywords Power system network · Single line to ground fault · Industrial load · Fault analysis · Typhoon HIL

R. Rajesh Kanna · R. Raja Singh (✉)
School of Electrical Engineering, Vellore Institute of Technology, Vellore, India

R. Raja Singh
Department of Energy and Power Electronics, Vellore Institute of Technology, Vellore, India

D. Arun Dominic
Department of Electrical and Electronics Engineering, NIT Karnataka, Surathkal, India

Nomenclature

E_x, E_y, E_z	Electromotive force per step during unsymmetrical fault
V_x, V_y, V_z	Phase voltages of three-phase transmission line
I_x, I_y, I_z	Phase current of three-phase transmission line
I_x^0, I_x^+, I_x^-	Zero, positive, negative sequence components of phase current
V_x^0, V_x^+, V_x^-	Zero, positive, negative sequence components of phase voltages
$Z_{KK}^0, Z_{KK}^+, Z_{KK}^-$	Zero, positive, negative sequence impedances
E_0, E_1, E_2	Zero, positive, negative sequence component of electromotive force
a	Sequence operator
L_{aa}, L_{bb}, L_{cc}	Self-inductance of three-phase transmission line
M_{ab}, M_{bc}, M_{ca}	Mutual inductance of three-phase transmission line
C_{ab}, C_{bc}, C_{ca}	Capacitance between two phases
C_{ea}, C_{eb}, C_{ec}	Capacitance between phase and ground
R, R_e	Resistance of transmission line and ground resistance

1 Introduction

Electrical power networks, primarily due to faults and incorrect processes, are often prone to disruptions. A fault is categorized as any interruption that affects the ordinary flow of power and disturbs the consistency, protection and superiority of the electricity supplied [1]. One of the most critical components of a power grid that links the generating station and the load centers is the transmission line. If the generating stations are so far from the load centers, then they operate over for a few hundred kilometers [2]. Unlike other parts of the power grid, the transmission line is mostly affected by fault because it is wide-open to the environment [3]. These are the main reasons to fault which arises on transmission line and are branches of tree falling on the transmission line, feeling of cross arm, measurement transformer blast, lightning, birds, animals and insulator material contamination [4]. It is possible to classify overhead transmission line faults as symmetrical as well as unsymmetrical faults. However, three-phase faults are the symmetrical fault in which all three phases are equally affected, sometimes it is also called as balanced fault. Unlike symmetrical faults, asymmetrical faults will not disturb all the three-phase lines equally [5]. Statistics indicate that all types of fault frequently occurred in the power system network. Among this, 85% of them are single line to the ground faults [6]. When a single line to ground fault takes place, some of the problems may occur: the voltage of the phases except faulted phase will get increased; an sporadic arc grounding may induce arc voltage spikes and feeder inaccessibility and easily lead to a create short circuit between phases [7]. To calculate the value of these fault currents as well as

fault megavolt-ampere, it is essential to execute fault analysis in the power system network [8]. According to the survey [9], the industrial loads are mostly influenced by the single phase to ground fault. The induction motor is definitely one of the most used electric machines in the world. The induction motor is widely adopted in many industries due to its reliability and robustness [10]. Due to the development of the Induction application, it is also used in various unusual facilities, such as heavy factory operations, engine running and ship impulsion. [11]. Here, for analyzing single line to ground fault in power system, 2.2 kW induction motor is used as a load in this category. Here this paper analyzes the impact of single line to ground fault on induction motor load under three different regions and different load values such as 0.25, 0.5 and 1 p.u. This paper approaches to Typhoon software in which 1 km/22 kV transmission model is simulated and single line to ground fault occurred by using fault tool box. After that, various effects on Induction motor load are shown such as current, voltage and power in terms of waveform.

The organization of this paper is as follows: The detailed mathematical modeling of single line to ground fault for the transmission system is presented in Sect. 2. Followed by, the impact of single line to ground fault on industrial load is realized using Typhoon HIL in Sect. 3. In Sect. 4, the estimation of transmission parameters such as R, L and C during the single line to ground fault is presented. Finally, the simulation results are obtained and discussed with three test cases in Sect. 5 and concluded in Sect. 6.

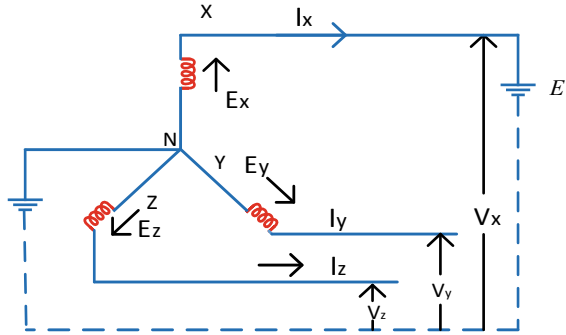
2 Mathematical Modeling of Single Line to Ground Fault

This section gives the detailed information about modeling of single line to ground fault. For this, the following conclusions can be created in the study of unsymmetrical faults [12]: (a) The generated electromotive force has only a positive sequence. (b) Load currents are ignored. (c) There is no fault impedance. (d) Phase 'x' should be recognized as the reference phase. Electromotive force per step is defined by E_x , E_y and E_z , in any case of unsymmetrical failure and phase voltages termed as V_x , V_y and V_z . The phase currents are defined as I_x , I_y and I_z . Consider neutral of a three-phase system must be earthed. An occurrence of a single line to ground fault on the phase 'x' of three-phase system is described in Fig. 1. It is clear from Fig. 1 that

$$V_x = 0 \text{ and } I_y = I_z = 0 \quad (1)$$

V_x is potential difference between phase 'x' and neutral. When single line to ground fault happens, the value of V_x will be zero. The zero, positive and negative sequence currents in the phase 'x' in terms of line currents are defined as I_x^0 , I_x^+ and I_x^- , respectively, which can be derived as in Eqs. (2)–(5),

Fig. 1 Single line to ground fault on three-phase system



$$I_x^0 = \frac{1}{3}(I_x + I_y + I_z) = \frac{1}{3}I_x \tag{2}$$

$$I_x^+ = \frac{1}{3}(I_x + aI_y + a^2I_z) = \frac{1}{3}I_x \tag{3}$$

$$I_x^- = \frac{1}{3}(I_x + a^2I_y + aI_z) = \frac{1}{3}I_x \tag{4}$$

$$I_x^0 = I_x^+ = I_x^- = \frac{1}{3}I_x \tag{5}$$

Here the operator ‘*a*’ is used to find the relationship between phase voltage and phase current. The positive, negative and zero sequence impedances of the generator are termed as Z_{KK}^+ , Z_{KK}^- and Z_{KK}^0 , respectively. Likewise the positive, negative and zero sequence components of phase current are I_x^+ , I_x^- and I_x^0 . From the closed loop *NXEN*, it is clear that the sequence currents create drops in voltage due to their corresponding sequence impedances, consequently, the electromotive force per step can be derived as in Eq. (6),

$$E_x = I_x^+ Z_{KK}^+ + I_x^- Z_{KK}^- + I_x^0 Z_{KK}^0 + V_x \tag{6}$$

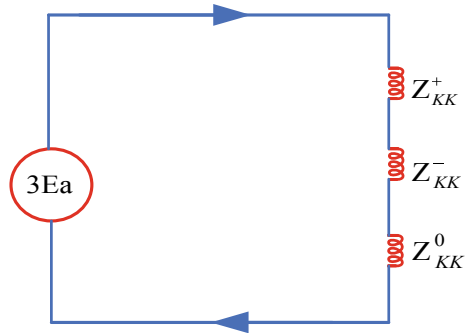
We know that

$$V_x = 0 \text{ and } I_x^+ = I_x^- = I_x^0 \tag{7}$$

From Eqs. (6) and (7), the zero sequence current is derived as in Eq. (8).

$$I_x^0 = \frac{E_x}{Z_{KK}^+ + Z_{KK}^- + Z_{KK}^0} \tag{8}$$

Fig. 2 Three-phase sequence impedance with imaginary generator voltage



With the help of creating the circuit which has consecutively connected the phase sequence impedances with an imaginary generator of voltage $3E_a$, fault current can be obtained which is described in Fig. 2. If the fault impedance is Z_f , then equation of fault current can be derived as in Eq. (9),

$$I_x = \frac{3E_x}{Z_{KK}^+ + Z_{KK}^- + Z_{KK}^0 + Z_f} \quad (9)$$

In case of the neutral is not connected to ground, then the impedance of the zero sequence component will be maximized and there is fault current flow since no path is there to flow of the fault current. The zero, positive and negative sequences of electromotive force system are E_0 , E_1 and E_2 , respectively. Since the produced electromotive force system is of positive sequence only, the sequential components of electromotive force in the phase 'x' are given in Eq. (10).

$$E_0 = 0, E_2 = 0 \text{ and } E_1 = E_x \quad (10)$$

The sequence components voltages at the fault for phase 'x' can be derived as in Eqs. (11)–(14)

$$V_x^+ = E_x - I_x^+ Z_{KK}^+ = E_x - \frac{E_x Z_{KK}^+}{Z_{KK}^+ + Z_{KK}^- + Z_{KK}^0} \quad (11)$$

$$V_x^- = \frac{Z_{KK}^- + Z_{KK}^0}{Z_{KK}^+ + Z_{KK}^- + Z_{KK}^0} E_x \quad (12)$$

$$V_x^0 = 0 - Z_{KK}^- I_x^- = \frac{-Z_{KK}^-}{Z_{KK}^+ + Z_{KK}^- + Z_{KK}^0} E_x \quad (13)$$

$$V_x^0 = 0 - I_x^0 Z_{KK}^0 = \frac{-Z_{KK}^0}{Z_{KK}^+ + Z_{KK}^- + Z_{KK}^0} E_x \quad (14)$$

From above equations, it can be concluded that $\vec{V}_x + \vec{V}_y + \vec{V}_z = 0$, because phase 'x' is connected to earth. The positive, negative and zero sequence components of phase voltages are V_x^+ , V_x^- and V_x^0 , respectively. Therefore, the phase voltages at fault can be written as in Eqs. (15)–(17).

$$V_x = V_x^+ + V_x^- + V_x^0 = 0 \quad (15)$$

$$V_y = V_x^0 + a^2 V_x^+ + a V_x^- \quad (16)$$

$$V_z = V_x^0 + a V_x^+ + a^2 V_x^- \quad (17)$$

3 Realization of Single Line to Ground Fault on Industrial Loads

This section provides the brief explanation about how to create the single line to ground fault on three-phase transmission line using Typhoon HIL and its responses. During a single line to ground fault in the transmission system, various impacts will be created on load section according to the distance of fault occurrence. For the implementation of the fault in the three-phase generation, transmission and distribution system, the Typhoon Virtual HIL software was used. Figure 3 shows the schematic representation of fault on three-phase power system.

A three-phase electrical power system consists of 11 kV, 50 Hz transmitting power from a three-phase source with 1500 kVA rating interconnected to induction motor load through a 1 km (1000 m) transmission line. The transmission line is divided into two 0.5 km (500 m) lines, connected between three-phase step-up and step-down transformers. In the proposed simulation, the three-phase transmission lines are used to interconnect three-phase source to load sector. The control system plan is presented in Fig. 5 which demonstrates the actions of the transmission line under live conditions to decide the resistance, inductance and capacitance of the transmission line [13]. The rating of various components had been carried out using a proper assumption and chosen with the help of below Eqs. (18)–(27). Properties of this model are the length of the line in kilometers, per length resistance and ground resistance in ohms/km, per length core and shield capacitance in farads/km, per length inductance and mutual inductance between phases in Henries/km and execution rate in seconds. A contactor switch takes into account the single line to ground fault which was only the switch that can connect any step to the ground.

Figure 4 gives the explanation about implementation of simulation in Typhoon HIL. At first, the data for three-phase system should be collected from library explorer

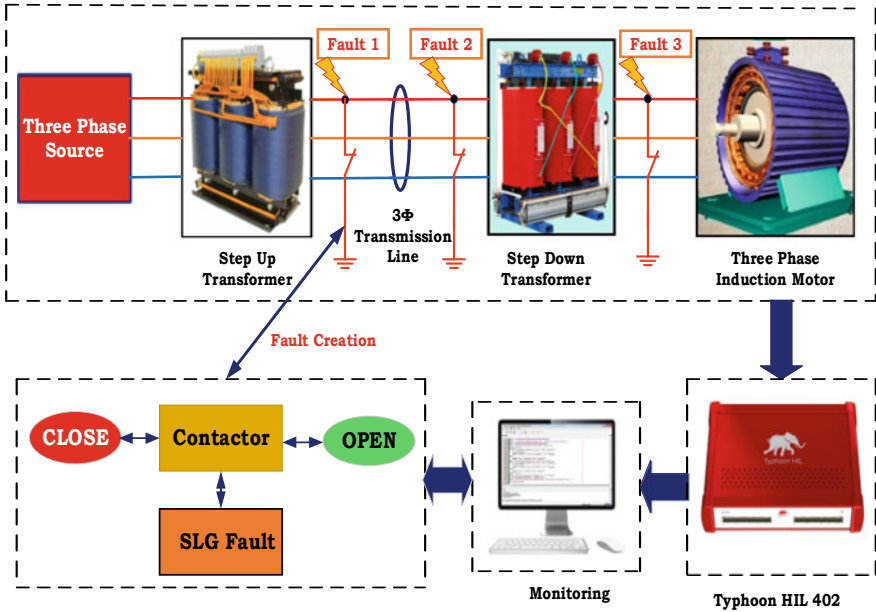


Fig. 3 Fault on three-phase transmission system

which is located in schematic editor and assembled as per Fig. 3. Then, in the model initialization script, write a Python program if it is required, once assembling of the three-phase system has been completed. Once the model file is created, it should be validated for upload in HIL SCADA. In case it is not validated, the assembled model and Python program have to be checked. If the model is validated, option virtual device should be chosen to carry out the output. Then different widgets can be dragged from SCADA panel to display various outputs. Once the widget creation is completed, the single line to ground fault should be created using close contactor in SCADA panel. For choosing various locations to create a single line to ground fault, again simulation model file has to be modified and compiled in the schematic editor. Finally, all the output of single line to ground fault for various locations has been compared and analyzed. The value of various electrical equipment is displayed in Table 1.

4 Transmission Line RLC Parameter Estimation

In this section, the estimation of RLC in three-phase transmission line has been explained. For a three-phase transmission line, the value of resistance, inductance

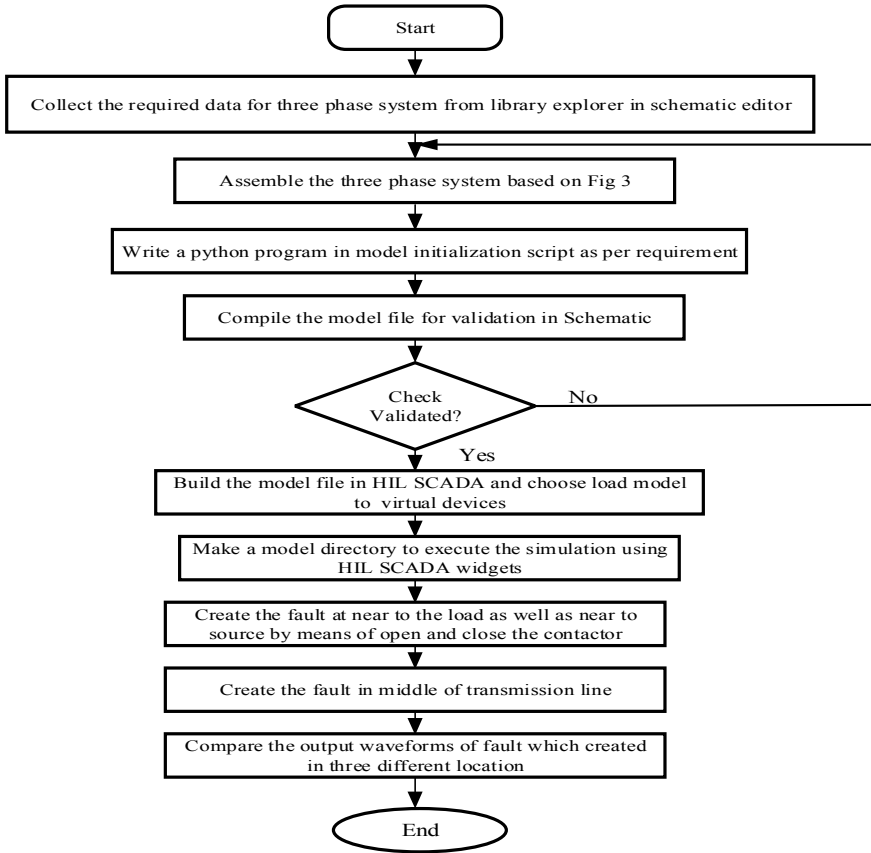


Fig. 4 Flowchart of Typhoon simulation execution

Table 1 Parameters for electrical instrument

Parameters	Three-phase source	Step-up transformer	Step-down transformer	Induction motor load
Rated power	1500 kVA	1500 kVA	1500 kVA	2.2 kW
Rated voltage	11 kV	11 kV/22 kV	22 kV/440 V	440 V
Rated current	–	1.3 kA/681 A	681 A/34 kA	4.5 A
Rated frequency	50 Hz	50 Hz	50 Hz	50 Hz

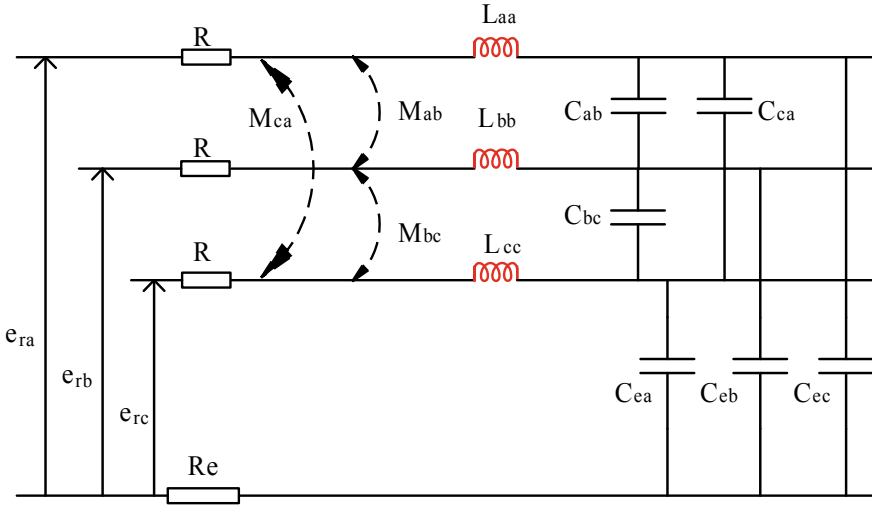


Fig. 5 Model of overhead three-phase transmission line

and capacitance value can be calculated using Eqs. (18)–(27). Figure 5 shows the basic structure and its availed parameters by which the equation has been made.

If the phase transmission line is balanced line, the inductance, capacitance and mutual inductance can be written as

$$L_{aa} = L_{bb} = L_{cc} = L_s; M_{ab} = M_{bc} = M_{ca} = M;$$

$$C_{ab} = C_{bc} = C_{ca} = C_c; C_{ea} = C_{eb} = C_{ec} = C_e.$$

Mathematical model of the above circuit can be expanded in the matrix form as follows:

$$-\frac{\partial}{\partial x}[e] = [L] \frac{\partial}{\partial t}[i] \tag{18}$$

$$-\frac{\partial}{\partial x}[i] = [C] \frac{\partial}{\partial t}[e] \tag{19}$$

where matrices [L] and [C] are

$$L = \begin{pmatrix} L_s & M & M \\ M & L_s & M \\ M & M & L_s \end{pmatrix} \tag{20}$$

$$C = \begin{pmatrix} 2C_c + C_e & -C_e & -C_e \\ -C_e & 2C_c + C_e & -C_e \\ -C_e & -C_e & 2C_c + C_e \end{pmatrix} \quad (21)$$

Model of the three-phase system is transformed into three decoupled single-phase line models, so-called modal lines, where transform matrix T is

$$T = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \quad (22)$$

Inverse transform matrix is

$$T^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & 0 \\ \frac{1}{3} & 0 & -\frac{1}{3} \end{pmatrix} \quad (23)$$

To transform matrices $[L]$ and $[C]$ into modal domain, we use transform matrix

$$[L_{\text{mode}}] = [T^{-1}][L][T] \quad (24)$$

$$[C_{\text{mode}}] = [T^{-1}][C][T] \quad (25)$$

Representation of the lossless three-phase power line can be expanded to model losses by adding series conductors' resistances and ground resistance.

$$R = \begin{pmatrix} R + R_e & R_e & R_e \\ R_e & R + R_e & R_e \\ R_e & R_e & R + R_e \end{pmatrix} \quad (26)$$

To transform $[R]$ into modal domain:

$$[R_{\text{mode}}] = [T^{-1}][R][T] \quad (27)$$

The values of R , L and C were then defined, based on system specificities, such as frequency of the system determined at 50 Hz. Algorithm of single line to ground fault creation implemented for three-phase power system is depicted in Fig. 5.

5 Simulation, Result and Discussion

The proposed three-phase system is simulated in Typhoon HIL. The three-phase generator is connected to induction motor load through step-up, step-down transformer and 1 km (1000 m) transmission line. The load is subjected to a single line to ground fault condition at three various distances. Two nodes had been created to measure the value of line current, line voltage, power and power factor, etc., at different levels. Here the concentration of measurement had made near to the load, middle of transmission line and near to source side. The simulation of three-phase system and SCADA panel output for simulation model is shown in Figs. 6 and 7.

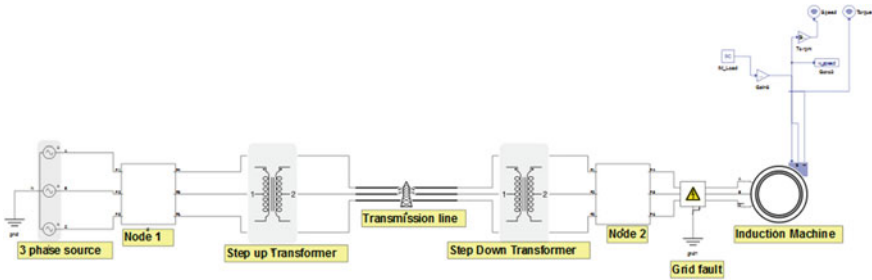


Fig. 6 Simulation model of three-phase system with grid fault

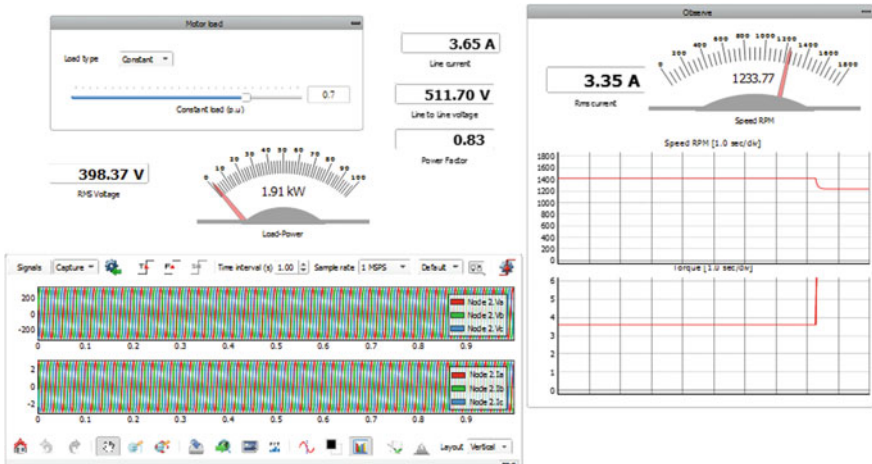


Fig. 7 SCADA panel output for simulation model of three-phase system

5.1 Test Case 1

In this section, the responses of system, when fault occurs on the transmission line, 1000 m from the load, are discussed. Simulated waveforms for changes in current and voltage when a fault occurs near to load are shown in Figs. 8, 9 and 10. When the fault occurs on phase R of the transmission line, at the fault point, the faulted phase current out of the power system and the remaining phases current reached zero. The faulted phase voltage will be equivalent to $V_R = Z_f I_R$. For bolted fault, the fault impedance is zero, and if an arcing fault occurs, the fault impedance will act as arc impedance. From Eqs. (8) and (9),

$$\text{Base current} = \frac{\text{MVA}_b}{\sqrt{3} * \text{KV}_b} \tag{28}$$

where

MVA_b = Base MVA rating of power system.

KV_b = Base kV rating of power system.

$$\text{Fault current in Amps} = I_x * \text{Base current} \tag{29}$$

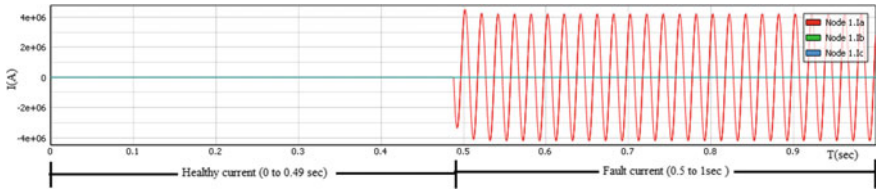


Fig. 8 Fault current at node 1 when fault occurs in beginning of transmission line

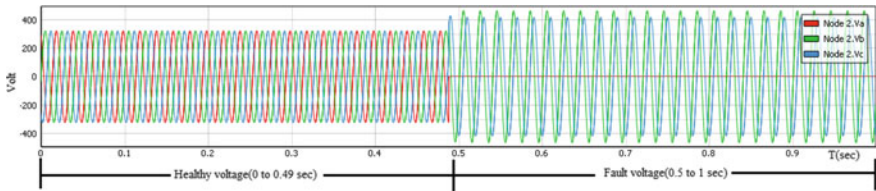


Fig. 9 Fault voltage at node 2 when fault occurs in beginning of transmission line

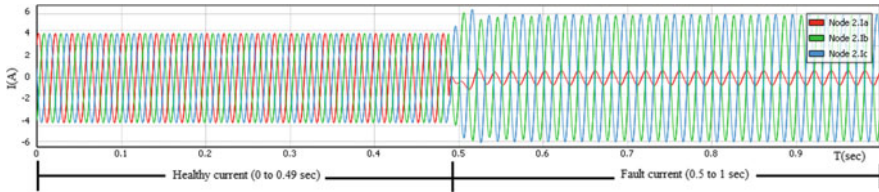


Fig. 10 Fault current at node 2 when fault occurs in beginning of transmission line

The positive sequence (Z_{KK}^+), negative sequence (Z_{KK}^-) and zero sequence impedance (Z_{KK}^0) have been assumed as $j0.002075$, $j0.002075$ and $j0.000572$ p.u. Once the impedances values are fixed, the fault current can be calculated by using Eqs. (28) and (29). Here as per the calculation, the value of fault current is 4.574×10^6 A, at source side when fault occurs near to the step-up transformer.

In Fig. 10, the faulted phase (R phase) is having some amount of current flow, even though that fell on the ground. The reason behind in current flow is single phasing effect of three-phase induction motor. Single phasing effect on the power system network would raise the current value in the left over two phases of a three-phase motor enlarged to 1.73 times of ordinary drawn current [14]. This is because the motor will try to maintain its rated output power until overload protecting equipment operates or until it burns out. Moreover, in a lightly loaded three-phase motor which is carrying 65 percent of normal full-load current, the phase current will theoretically increase by 1.73 times under secondary single phasing. Afterward, the motor current drawn will be around 1.12 times of full-load current as shown in nameplate details of the motor.

5.2 Test Case 2

In Sect. 5.2, the responses of system, when the fault occurs in middle of transmission line, 500 m from the load are discussed. Simulated waveforms for changes in current and voltage when a fault occurs in the center point of transmission line are shown in Figs. 11, 12 and 13. When single line to ground fault occurs in central of transmission

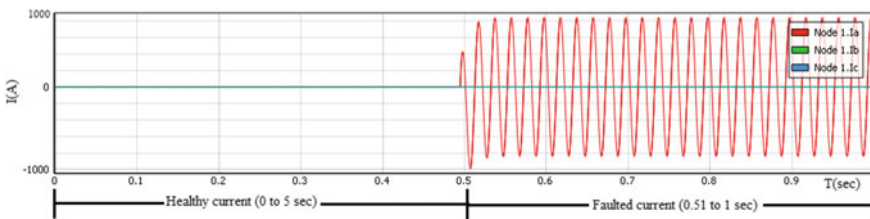


Fig. 11 Fault current at node 1 when fault arises central of transmission line

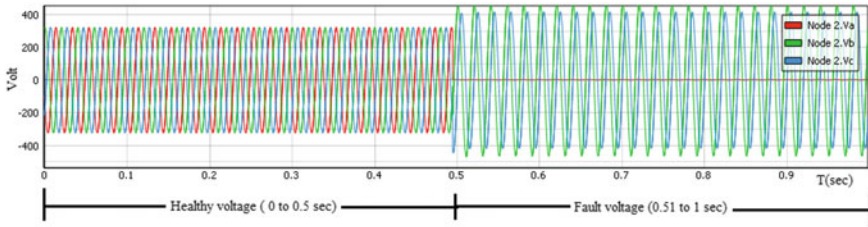


Fig. 12 Fault voltage at node 2 when fault arises central of transmission line

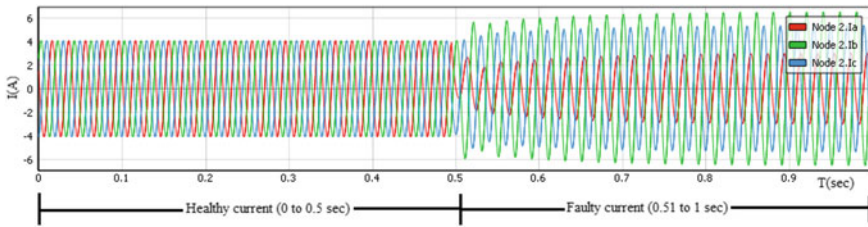


Fig. 13 Fault current at node 2 when fault arises central of transmission line

line, the effect on line current and voltage of load is a little more compared with fault occurs in beginning of transmission line.

5.3 Test Case 3

In Sect. 5.3, the responses of system, when the fault occurs in middle of transmission line, 10 m from the load are discussed. Simulated waveforms for changes in current and voltage when a fault occurs near to the induction motor load are shown in Figs. 14, 15 and 16. When single line to ground fault occurs in the middle of transmission line, the impact on induction motor load is very high compared with the other two cases.

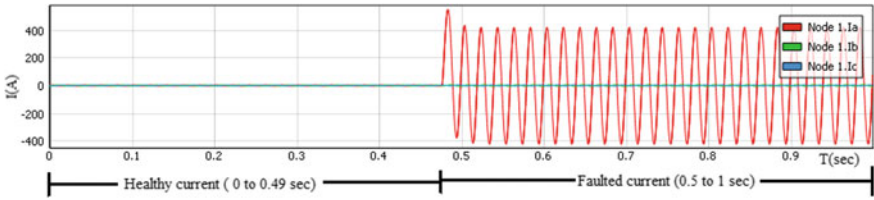


Fig. 14 Fault current at node 1 when fault occurs near to the load

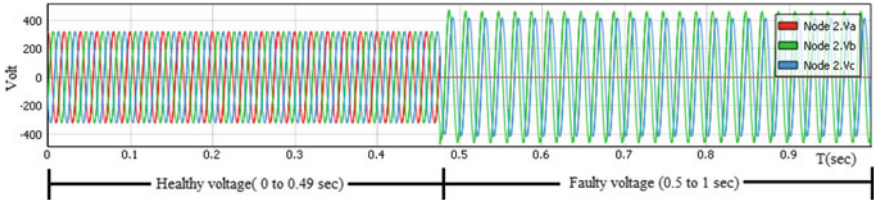


Fig. 15 Fault voltage at node 2 when fault occurs near to the load

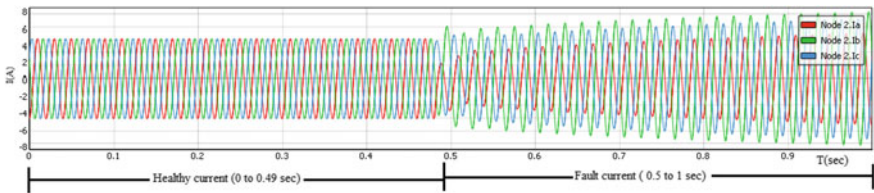


Fig. 16 Fault current at node 2 when fault occurs near to the load

6 Conclusions

The impact of single line to ground fault on industrial loads is analyzed, and the performance of the system according to the location of the fault is studied using Typhoon hardware in loop software. The paper presents the detailed simulation and execution procedure for implementing the proposed system. In the Typhoon HIL schematic editor, the three-phase power system with an industrial motor load is simulated, and the fault is created in three different locations such as 1000, 500 and 10 m. The impact of single line to ground fault on industrial loads under different loading conditions is analyzed and given in Table 2. It is perceivable from the analysis that the fault currents are more when it is created near to the load, and also it is evident that the loading is not having much impact.

Table 2 Fault current analysis at node 2 for various loads and distances

Load (p.u)	Variation of fault current from nominal value in percentage			
	Nominal current (A)	1000 m from load (%)	500 m from load (%)	10 m from load (%)
1	4.5	22*	42*	73*
0.5	2.2	24*	48*	72*
0.25	1.3	16*	37*	69*

*Increased

References

1. R.O. Ruback, V. Menengoy, S. Medeiros, Electrical power and energy systems a new method for analyzing unsymmetrical faults under data uncertainties. *Electr. Power Energy Syst.* **102**(April), 38–51 (2018). <https://doi.org/10.1016/j.ijepes.2018.04.022>
2. L. Shilong, T. Yufei, L. Xiaopeng, Z. Huajie, F. Shilin, Fault line selection of single phase grounding fault in small-current ground system based on reactive current, pp. 138–143 (2019)
3. S. Karekar, S. Kumar, A. K. Sahu, Modeling and simulation of overhead transmission line faults detection and comparison **March**, 196–200 (2016)
4. M.H. Idris, M.W. Mustafa, Y. Yatim, Effective two-terminal single line to ground fault location algorithm **June**, 6–7 (2012)
5. T.Y. Sing, S. Emran, B. Siraj, R. Raguraman, P.N. Marimuthu, K. Nithiyananthan, *Int. J. Adv. Appl. Sci. Cosine Similarity Cluster Analys. Mod. Based Effect. Pow. Syst. Fault Identif.* **4**(1) (2017)
6. S. Bagchi, S. Goswami, B. Ghosh, M. Dutta, R. Bhaduri, Symmetrical and asymmetrical fault analysis of transmission line with circuit breaker operation, pp. 343–347 (2019)
7. X. Wang, J. Gao, X. Wei, Z. Zeng, Y. Wei, M. Kheshti, Single line to ground fault detection in a non-effectively grounded distribution network. *IEEE Trans. Power Deliv.* **MM**, 1 (2018). <https://doi.org/10.1109/tpwrld.2018.2873017>
8. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON) Held at Deenbandhu Chhotu Ram University of Science and Technology, SONEPAT, India on FEB 28 - March 1 2020*
9. R. Raja Singh, C. Thanga Raj, R. Palka, V. Indragandhi, M. Wardach, P. Paplicki, Energy optimal intelligent switching mechanism for induction motors with time varying load. *IOP Conf. Ser. Mater. Sci. Eng.*, **906**(1) (2020). <https://doi.org/10.1088/1757-899x/906/1/012017>
10. H.A.N. Zeng, P. Yang, H. Cheng, J. Xin, Research on single-phase to ground fault simulation base on a new type neutral point flexible grounding mode. *IEEE Access* **7**, 82563–82570 (2019). <https://doi.org/10.1109/ACCESS.2019.2922361>
11. A. Birajdar, S. Tajane, Modelling and simulation of transmission line to detect single line to ground fault location, pp. 1–4 (2016)
12. V.K. Mehta, R. Mehta, *Principles of power system* (2016)
13. www.typhoon-hil.com

14. I.Z. Giceva, V.J. Sarac, S.A. Gelev, V.T. Cingoski, Single phasing of three phase induction motors under various load conditions, in *2018 23rd International Scientific-Professional Conference on Information Technology (IT)*, Zabljak, 2018, pp. 1–4. <https://doi.org/10.1109/spit.2018.8350459>

Performance Improvement for PMSG Tidal Power Conversion System with Fuzzy Gain Supervisor Passivity-Based Current Control



Youcef Belkhier, Abdelyazid Achour, Rabindra Nath Shaw, and Ankush Ghosh

Abstract The power captured by a tidal conversion system depends highly on the applied control strategies. In fact, nonlinear properties of the generator, parameter uncertainties, and external disturbances make the controller design a challenging problem. This paper contributes with the novel fuzzy gain supervisor passivity-based control (FGSPBC) method that allows the PBC to be faster, combined with a PI controller where its gains are adopted by the FGSPBC, applied to the PMSG based variable tidal turbine with grid connection via back-to-back converter to bring the PMSG to work at an optimal point while ensuring stability, fast convergence of the conversion system, and performance improvement. The aims of this work are that the regulation of the DC voltage and the reactive power at their respective values, whatever the disturbances related to the PMSG. Numerical investigation under MATLAB/Simulink addresses the effectiveness, stability, and fast convergence of the studied system.

Keywords Adaptive control · Parameter changes · Passive control · Tidal conversion system

1 Introduction

Recently, various turbine-based PMSGs have been developed to extract tidal energy [1]. The most widely used technology to convert marine current energy consists of a tidal generator, a PMSG, power converter, and load, due to its advantages such

Y. Belkhier (✉) · A. Achour

Laboratoire de Technologie Industrielle et de l'Information (LTII), Faculté de Technologie, Université de Bejaia, 06000 Bejaia, Algeria
e-mail: youcef.belkhier@univ-bejaia.dz

R. N. Shaw

Department of Electronics and Communication Engineering, Galgotias University, Greater Noida, India

A. Ghosh

School of Engineering and Applied Sciences, The Neotia University, Sarisha, West Bengal, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_6

81

as predictability, low cost, and clean energy. However, the controller design of this kind of conversion system-based PMSG is still a challenging work due to the PMSG issues such as time-varying parameters and external disturbances [2, 3].

This paper contributes with the novel fuzzy gain supervisor passivity-based control (FGSPBC) method that allows the PBC to be faster, combined with a PI controller where its gains are adopted by the FGSPBC, applied to the PMSG based variable tidal turbine with grid connection via back-to-back converter to bring the PMSG to work at an optimal point while ensuring stability, fast convergence of the conversion system, and performance improvement. The maximum power extraction from the wind turbine, with taking into account its entire dynamic when synthesizing the controller, represents the main motivation of the present work. The present controller performance is compared with those of the conventional PI and the present strategy without the fuzzy supervisor (PI with fixed gains) [4].

The present paper is organized by the present form: in Sect. 2, the system description is established. Section 3 deals with the proposed strategy computation. Concerning Sect. 4, the grid-side converter (GSC) PI controller is formulated. Section 5 presents the numerical validation of the presented control strategy. Finally, main conclusions are presented in Sect. 6.

2 Tidal Power and PMSG Modeling

2.1 Tidal Power

The power which can be extracted with the turbine from the tidal energy and its related output torque T_m , are expressed as (see Fig. 1) [1–3]:

$$P_m = \frac{1}{2} \rho C_p(\beta, \lambda) A v^3 \tag{1}$$

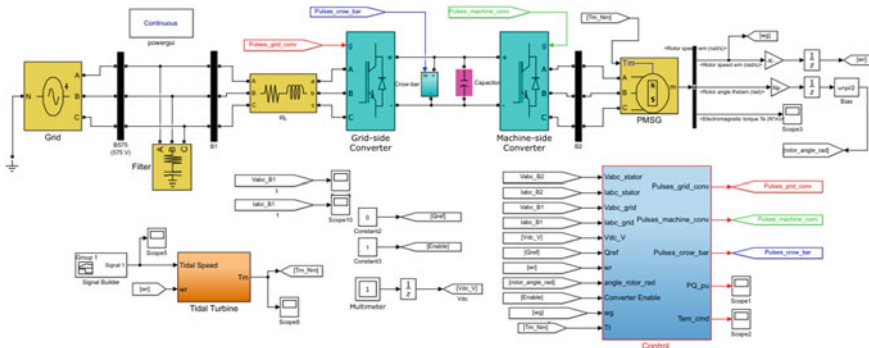


Fig. 1 Studied system with MATLAB/Simulink

$$C_p(\beta, \lambda) = \frac{1}{2} \left(\frac{116}{\lambda_i} - 0.4\beta - 5 \right) e^{-\left(\frac{21}{\lambda_i}\right)} \quad (2)$$

$$\lambda_i^{-1} = (\lambda + 0.08\beta)^{-1} - 0.035(1 + \beta^3)^{-1} \quad (3)$$

$$T_m = \frac{P_m}{\omega_m} \quad (4)$$

where $\lambda = \frac{\omega_m R}{v}$ denotes the ratio of the tip speed, A denotes the area of the blades, C_p denotes the coefficient of the power, β denotes the angle of the pitch, ρ denotes the tidal density, v denotes the speed of the tidal, ω_t represents the speed of the turbine, and R denotes the blades radius.

2.2 PMSG Modelling

The PMSG model is given by [3, 4] in dq reference frame as follow:

$$v_{dq} = R_{dq} i_{dq} + L_{dq} \dot{i}_{dq} + p\omega_m \mathfrak{S}(L_{dq} i_{dq} + \psi_f) \quad (5)$$

$$J\dot{\omega}_m = T_m - T_e - f_{fv}\omega_m \quad (6)$$

$$T_e = \frac{3}{2} p \psi_{dq} \mathfrak{S} i_{dq} \quad (7)$$

where $R_{dq} = \begin{bmatrix} R_s & 0 \\ 0 & R_s \end{bmatrix}$ denotes the stator resistances matrix, f_{fv} represents the coefficient of the viscous friction, $\psi_f = \begin{bmatrix} \phi_f \\ 0 \end{bmatrix}$ denotes the flux linkages, T_e represents the electromagnetic torque, $L_{dq} = \begin{bmatrix} L_d & 0 \\ 0 & L_q \end{bmatrix}$ denotes the stator inductions matrix, $v_{dq} = \begin{bmatrix} v_d \\ v_q \end{bmatrix}$ denotes the vector of the stator voltages, J denotes the inertia moment, and $\mathfrak{S} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $i_{dq} = \begin{bmatrix} i_d \\ i_q \end{bmatrix}$ denotes the vector of the stator currents.

3 Proposed Method Design

Making the system passive is the principal aim of the proposed method. This is possible by introducing a damping term and reshaping its energy. The application of the adaptive fuzzy supervisory high-order passivity-based combines nonlinear observer and voltage control developed in this paper needs numerous steps: The controller design process is shown in Fig. 2, in which, two main parts can be distinguished: firstly, the reference dynamic that consists in the desired current, computed through the desired torque, calculated by the proposed fuzzy supervisory PI controller. Then, the voltage controller is computed by the passivity-based current controller.

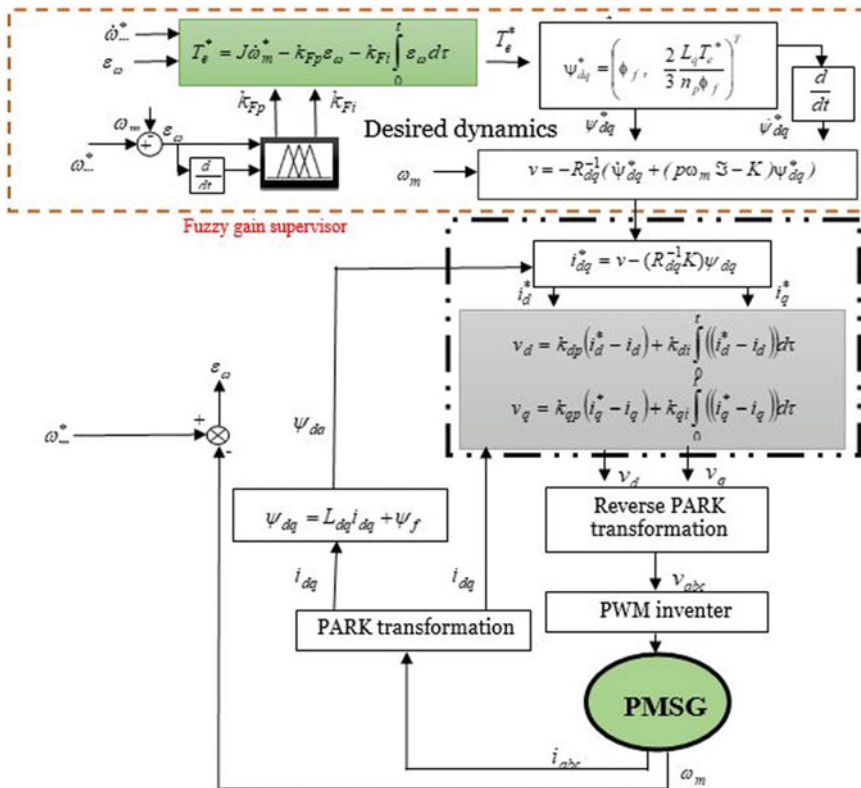


Fig. 2 Proposed control strategy block diagram

3.1 Passivity-Based Current Controller Computation

The desired current vector i_{dq}^* which is considered as the PMSG control inputs is defined as follow:

$$i_{dq}^* = \begin{bmatrix} i_d^* \\ i_q^* \end{bmatrix} \quad (8)$$

and the controller output is the vector of the voltage which is v_{dq} computed as follow [4]:

$$v_d = k_{dp}\varepsilon_{di} + k_{di} \int_0^t \varepsilon_{di} d\tau \quad (9)$$

$$v_q = k_{qp}\varepsilon_{qi} + k_{qi} \int_0^t \varepsilon_{qi} d\tau \quad (10)$$

where $k_{dp} > 0$, $k_{di} > 0$, $k_{qp} > 0$ and $k_{qi} > 0$. $\varepsilon_i = \begin{bmatrix} \varepsilon_{di} \\ \varepsilon_{qi} \end{bmatrix} = \begin{bmatrix} i_d^* - i_d \\ i_q^* - i_q \end{bmatrix}$ is the currents error vector. In order to track its convergence, the dynamics of the current vector are simplified, and their reference values are considered as the control inputs of the PMSG when we assume that the PI current loops work satisfactorily by the proper choices of the positive gains. Then, we formulate the following desired dynamic of the PMSG:

$$\dot{\psi}_{dq} + p\omega_m \Im \psi_{dq} = -R_{dq} i_{dq}^* \quad (11)$$

$$J\dot{\omega}_m = T_m - T_e^* - f_{fv}\omega_m \quad (12)$$

$$T_e = \frac{3}{2} p \psi_{dq} \Im i_{dq}^* \quad (13)$$

where T_e^* represents the desired torque and ψ_{dq} denotes the flux linkages expressed by [4]:

$$\psi_{dq} = \begin{bmatrix} \psi_d \\ \psi_q \end{bmatrix} = \begin{bmatrix} L_d i_d + \phi_f \\ L_q i_q \end{bmatrix} \quad (14)$$

By taking the desired vector of the flux linkage as $\psi_{dq}^* = \begin{bmatrix} \psi_d^* \\ \psi_q^* \end{bmatrix}$, its tracking error vector as $e_f = \begin{bmatrix} e_{fd} \\ e_{fq} \end{bmatrix} = \psi_{dq} - \psi_{dq}^*$, and replacing e_f in (11), the expression of e_f is deduced. Controller input $i_{dq}^* = \begin{bmatrix} i_d^* \\ i_q^* \end{bmatrix}$ is deduced by using the Lyapunov theory and the defined function $V(e_f) = 0.5e_f^T e_f$. Then, it yields the following expression:

$$i_{dq}^* = -\frac{1}{R_{dq}}(K e_f - (\dot{\psi}_{dq}^* + p\omega_m \Im \psi_{dq}^*)) \quad (15)$$

where $K = \begin{bmatrix} K_1 & 0 \\ 0 & K_2 \end{bmatrix}$, $K_1 > 0$ and $K_2 > 0$.

The PMSG operates at optimal torque only if i_d is fixed to zero. Under this condition, the flux ψ_d is reduced to the flux ϕ_f created by the permanent magnet as follow:

$$\psi_d^* = \phi_f \quad (16)$$

By using (13) and (14), the desired torque is computed as given below:

$$T_e^* = -\frac{3}{2}p\psi_{dq}^{*T} \Im(\psi_{dq}^* - \psi_f) \quad (17)$$

From (16) and (17), following expression of the desired flux along q -axis is deduced:

$$\psi_q^* = \frac{2}{3} \frac{L_q}{p\phi_f} T_e^* \quad (18)$$

3.2 Fuzzy Gains Supervisor-PI Control of Desired Torque

The desired torque T_e^* deduced from (13) that yields:

$$T_e^* = J\dot{\omega}_m^* + T_m + f_{fv}\varepsilon_\omega \quad (19)$$

where $\varepsilon_\omega = \omega_m^* - \omega_m$ represents the speed error and ω_m^* denotes the turbine speed. From (19), two problems are deduced: the parameters (J/f_{fv}) restrict its convergence, and it is an open loop. To address this issue, the authors in removed the term (f_{fv}) , and PI term is adopted as expressed below:

$$T_e^* = J\dot{\omega}_m^* - k_{Fp}\varepsilon_\omega - k_{Fi} \int_0^t \varepsilon_\omega d\tau \tag{20}$$

where $k_{Fp} > 0$ and $k_{Fi} > 0$. However, PI controls are generally not robust for controlling systems with parameter changes and uncertainties, due to the fixed gains, such as discussed in. Thus, a fuzzy gain supervisor is introduced, to ensure fast convergence of ε_ω . The fuzzy supervisor is used to adjust the PI gains k_{Fp} and k_{Fi} , and thus, solve the problem caused by imprecise parameters, where the inputs are chosen, respectively, as ε_ω and its derivative $\Delta\varepsilon_\omega$. The fuzzy supervisor diagram is shown in Figs. 3, 4, 5, and 6.

The advantage of this method is that the number of parameters of the membership functions is significantly reduced. The linguistic variables corresponding to the inputs–outputs of the fuzzy gain scheduling are given in Tables 1 and 2 as in. Then, the design of the controller law with the proposed combined fuzzy gain supervisory-PI is shown in Fig. 2.

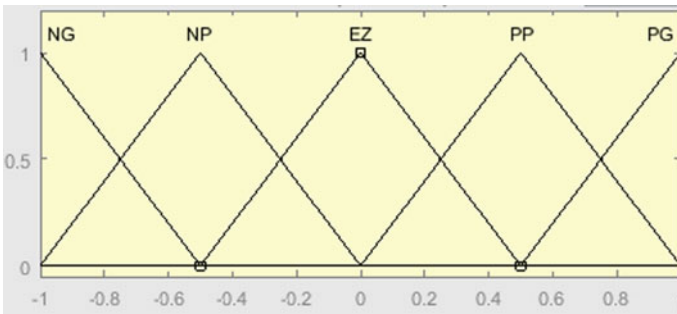


Fig. 3 Membership function of inputs ε_ω and $\Delta\varepsilon_\omega$

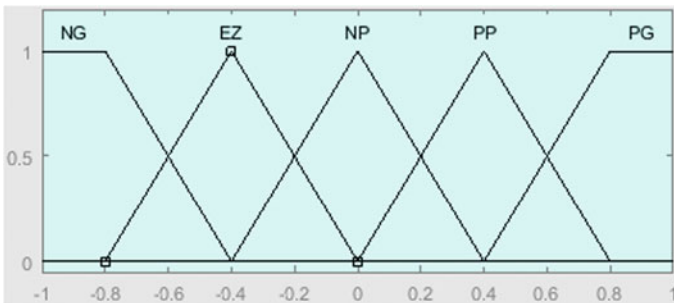


Fig. 4 Membership function of outputs k_{Fp} and k_{Fi}

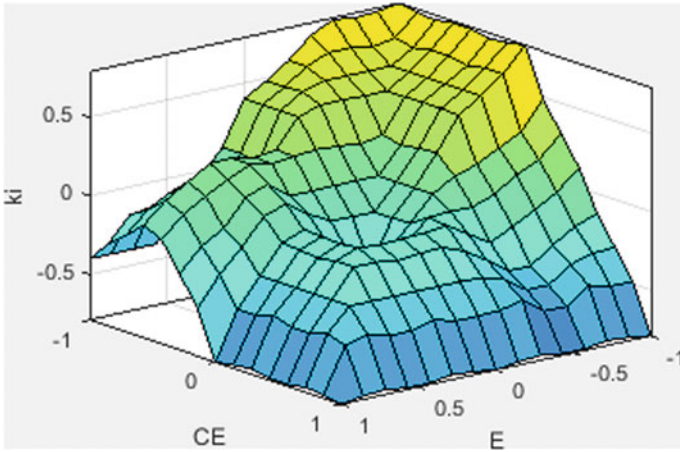


Fig. 5 k_{Fi} fuzzy surface

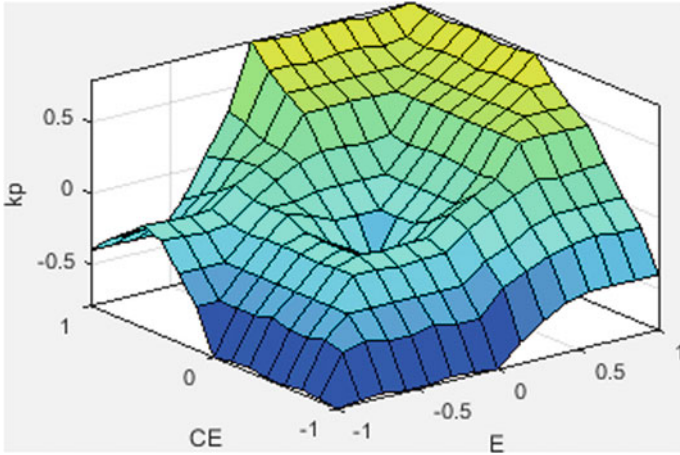


Fig. 6 k_{Fp} fuzzy surface

Table 1 Fuzzy rules of ε_ω and $\Delta\varepsilon_\omega$

ε_ω	$\Delta\varepsilon_\omega$				
	NB	NS	Z	PS	PB
NB	NB	NB	NB	NS	Z
NS	NB	NS	NS	NS	Z
Z	NB	NS	Z	PS	PB
PS	Z	PS	PS	PS	PB
PB	Z	PS	PB	PB	PB

Table 2 Fuzz rules of k_{Fp} and k_{Fi}

k_{Fp}	k_{Fi}				
	NB	NS	Z	PS	PB
NB	PB	PB	PB	NS	NB
NS	PB	PS	PS	Z	NB
Z	PS	PS	Z	NS	NB
PS	Z	PS	NS	NS	NB
PB	Z	NS	NB	NB	NB

4 Grid-Side PI Controller

The schema of the classical PI strategy is used for the GSC. Its principal aim is to transmit to the grid the active power and regulate the DC-bus voltage. The GSC mathematical description is given as below:

$$\begin{bmatrix} V_{id} \\ V_{iq} \end{bmatrix} = \begin{bmatrix} L_f \dot{i}_{df} - \omega L_f i_{qf} \\ L_f \dot{i}_{qf} - \omega L_f i_{df} \end{bmatrix} + R_f \begin{bmatrix} i_{df} \\ i_{qf} \end{bmatrix} + \begin{bmatrix} V_{gd} \\ V_{gq} \end{bmatrix} \quad (21)$$

$$C \dot{V}_{dc} = \frac{3}{2} \frac{v_{gd}}{V_{dc}} i_{df} + i_{dc} \quad (22)$$

where i_{df} and i_{qf} denote the grid currents, ω represents the grid angular frequency, V_{id} , V_{iq} denote the inverter voltages, V_{gd} , V_{gq} denotes the grid voltages, L_f represents the filter inductance, R_f denotes the filter resistance, V_{dc} represents the voltage of the DC-link, C denotes the capacitor of the DC-link, and i_{dc} represents DC current. Reactive power Q_g and active power P_g are expressed by:

$$\begin{cases} P_g = \frac{3}{2} v_{gd} i_{df} \\ Q_g = \frac{3}{2} v_{gd} i_{qf} \end{cases} \quad (23)$$

5 Simulation Results

In this part, extensive numerical investigation by using MATLAB/Simulink shown by Fig. 1 is performed to show the feasibility of the proposed strategy. The simulated conversion system is based on 1.5 MW rated power. The closed-loop parameter values are given in Table 3. The gains of the PBC PI controller are $k_{dp} = k_{qp} = 99$, $k_{di} = k_{qi} = 200$, and $K_1 = K_2 = 100$. The fixed gains of the desired torque for the proposed control without the fuzzy gain supervisor are $k_{Fp} = 5$ and $k_{Fi} = 1000$. The gains of the DC-link are $k_{dcp} = 100$, $k_{dci} = 8000$, and the gains of the currents PI

Table 3 Parameters of the system

PMSG parameter	Value
Tidal density (ρ)	1024 kg/m ²
DC-link capacitor (C)	2.9 F
Pole pairs number (p)	48
Stator inductance (L_{dq})	0.3 mH
Tidal turbine radius (R)	10 m
Grid-filter resistance (R_f)	0.3 pu
Stator resistance (R_s)	0.006 Ω
Grid-filter inductance (L_f)	0.3 pu
DC-link voltage (V_{dc})	1150 V
Flux linkage (ϕ_f)	1.48 Wb
Total inertia (J)	35,000 kg m

control are $k_{cp} = 9, k_{ci} = 200$. The proposed strategy is compared to the passivity-based current control (PBCC) namely (proposed without supervisor) proposed in [4] and the conventional PI method.

The tidal speed dynamic is shown in Fig. 7. Figure 8 presents the response of the electromagnetic torque, as shown the proposed control presents fast convergence and

Fig. 7 Tidal velocity

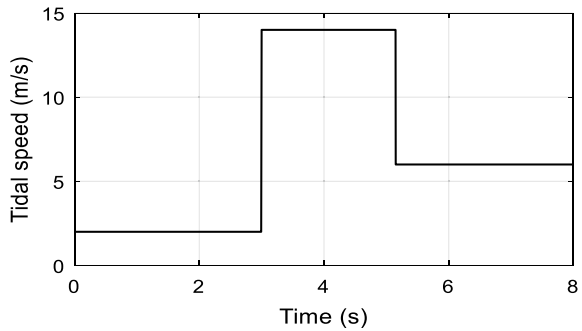
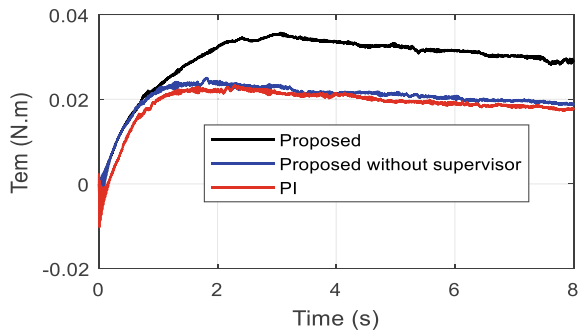


Fig. 8 Response of the electromagnetic torque



a higher torque than the PBCC and the conventional PI control. Figure 9 shows that the DC-link voltage quickly tracks its reference value 1150 V, without overshoot for the proposed strategy unlike the PBCC and the PI strategy. Figures 10 and 11 show that the reactive power very well kept at its zero-reference for all the control methods. However, as shown, the FGSPBCC presents a fast convergence than the PBCC and the PI. In Fig. 12, it can be deduced that the active power is the only one transmitted to the grid for all the control strategies with largely more power transmitted for the proposed one. In Fig. 13, we can see a perfect sinusoidal voltage absorption by the grid. The FGSPBCC shows a higher torque, a best power generation, than the conventional methods.

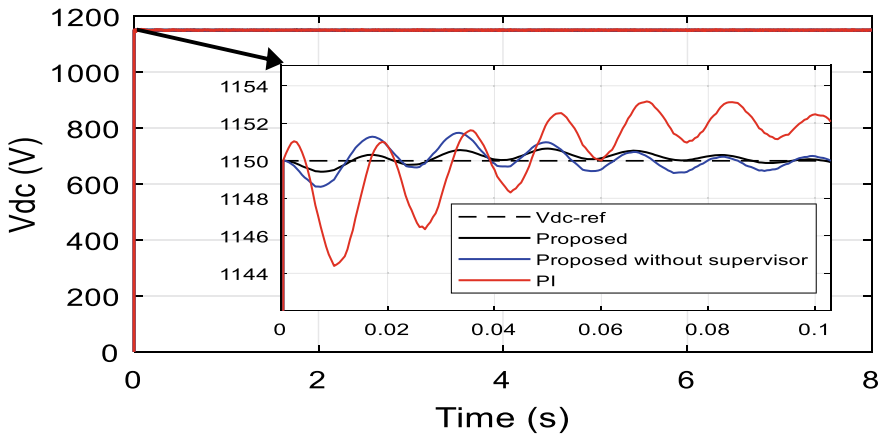


Fig. 9 Response of the DC-link voltage

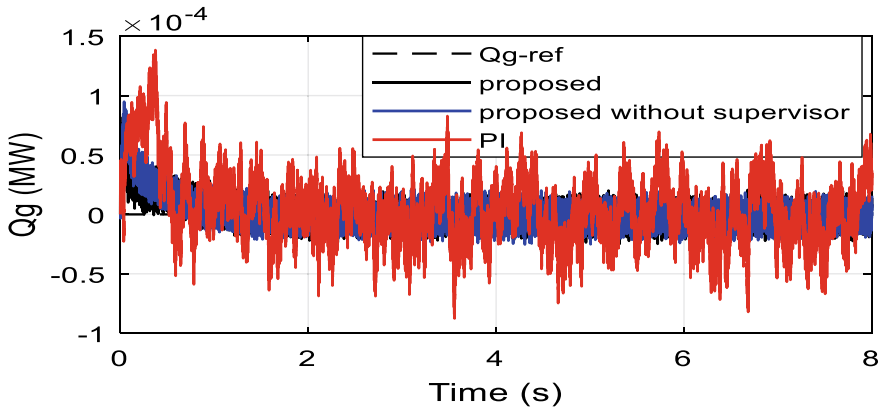


Fig. 10 Response of the reactive power

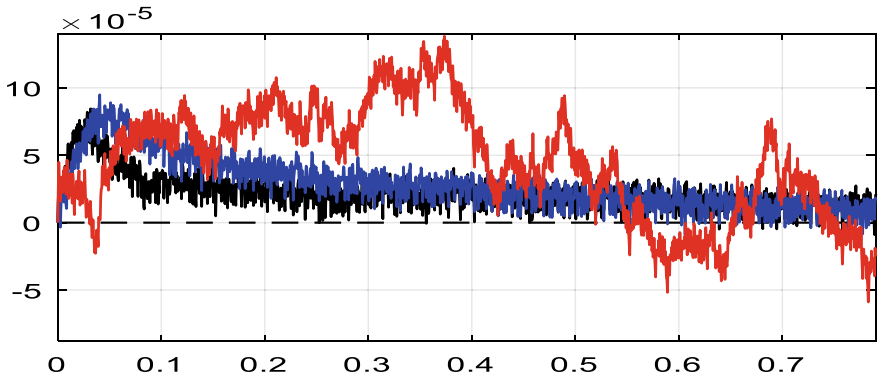


Fig. 11 Zoom on reactive power

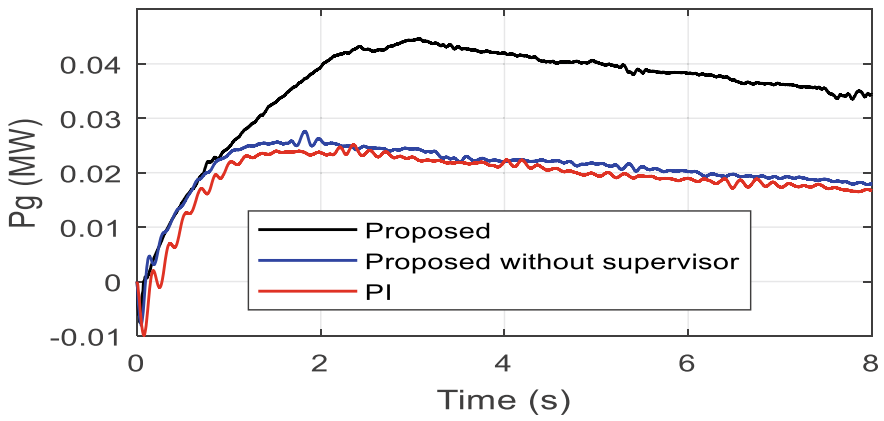


Fig. 12 Active power

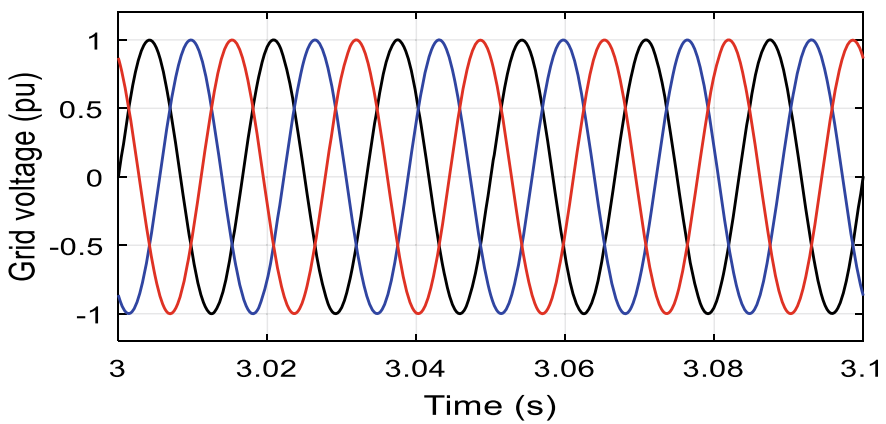


Fig. 13 Grid-injected voltage

6 Conclusion

In this study, a novel fuzzy supervisor passivity-based current controller for a tidal power system with PMSG is developed for overall improvement of the performances. The simulation results are performed by the way of MATLAB/Simulink environment which shows the effectiveness of the studied closed loop. The proposed strategy presents the fast tracking of the maximum wind power, the reactive power generated, and the DC-link which are well kept at their set values. All drawbacks of the presented controller are resolved, and the objectives achieved. Higher efficiency compared to the conventional PI method are provided by the proposed controller.

References

1. Y. Belkhier, A.Y. Achour, R.N. Shaw, Fuzzy passivity-based voltage controller strategy of grid-connected PMSG-based wind renewable energy system, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India (2020), pp. 210–214. <https://doi.org/10.1109/ICCCA49541.2020.9250838>
2. P. Qian, B. Feng, H. Liu, X. Tian, Y. Si, D. Zhang, Review on configuration and control methods of tidal current turbines. *Renew. Sust. Energ. Rev.* **108**, 125–139 (2019)
3. M.O.F. Diallo, S. Youssef, H. Gualous, M.B. Camara, B. Dakyo, Permanent magnet synchronous generator for tidal turbine application in Raz Blanchard—modeling and control strategy, in *2014 16th International Power Electronics and Motion Control Conference and Exposition*, IEEE, Antalya, Turkey (2014), pp. 377–381
4. A.Y. Achour, B. Mendil, S. Bacha, I. Munteanu, Passivity-based current controller design for a permanent-magnet synchronous motor. *ISA Trans.* **48**(3), 336–346 (2009)
5. Y. Belkhier, A.Y. Achour, Fuzzy passivity-based linear feedback current controller approach for PMSG-based tidal turbine. *Ocean Eng.* **218**, 108156 (2020). <https://doi.org/10.1016/j.oceaneng.2020.108156>
6. Y. Belkhier, A.Y. Achour, Passivity-based current control strategy for PMSG wind turbine, in *2019 IEEE 1st International Conference on Sustainable Renewable Energy Systems and Applications (ICSRESA)*, Tebessa, Algeria (2019), pp. 1–4
7. Z. Zhou, B.S. Elghali, M.E.H. Benbouzid, Y. Amirat, E. Elbouchikhi, G. Feld, Tidal stream turbine control: an active disturbance rejection control approach. *Ocean Eng.* **202**, 107190 (2020)
8. K. Ghafari, I. Garrido, A.J. Garrido, S. Bouallègue, J. Haggège, Fuzzy gain scheduling of a rotational speed control for tidal stream generator, in *2018 International Symposium on Power Electronics, Electrical Drives, Automation and Motion*, Amalfi, Italy (2018), pp. 1271–1277
9. Y. Krim, D. Abbes, S. Krim, M.F. Mimouni, Intelligent droop control and power management of active generator for ancillary services under grid instability using fuzzy logic technology. *Control Eng. Pract.* **81**, 215–230 (2018)
10. Y. Belkhier, A.Y. Achour, An intelligent passivity-based back stepping approach for optimal control for grid-connecting permanent magnet synchronous generator-based tidal conversion system. *Int. J. Energy Res.* 1–16 (2020). <https://doi.org/10.1002/er.6171>

Harmonics Minimization in Multilevel Inverter by Continuous Mode ACO Technique



Salman Ahmad, Atif Iqbal, Imtiaz Ashraf, and Zahoor Ahmad Ganie

Abstract Low switching frequency operation of multilevel inverter enables it for applications in high power processing needs. The stepped waveform synthesized from multilevel inverter after Fourier series analysis will result in system of nonlinear equations to selectively minimize certain lower-order harmonics from it. These equations are transcendental and exhibits solutions or no solution in certain range of modulation index. In order to obtain continuous solution in entire modulation index range optimization techniques based on meta-heuristic approach is used. In order to achieve better quality waveform at output and selectively removing certain low-order harmonics an ant colony optimization-based technique in continuous mode implementation is proposed in this paper. The simulation and experimental results are provided to confirm the correctness and effectiveness of the method.

Keywords Ant colony optimization · Pulse width modulation · Multilevel inverter

1 Introduction

Voltage source inverters (VSI) are extensively used in many industrial applications such as in variable speed electric drives, STATCOM, active filters, HVDC, flexible AC transmission, pump storage plants, and solar photovoltaic system of renewable energy harnessing [1, 2]. Earlier only two-level VSI most popularly used for these applications produces a bipolar waveform at the output. But since the power processing need increased, the need to develop high power processing converters, and efficient modulation and control techniques become an active area of research. Therefore, in this direction, several multi-phase topologies and series parallel combination

S. Ahmad (✉) · Z. A. Ganie

Islamic University of Science and Technology, Awantipora, J&K 192122, India

A. Iqbal

Department of electrical engineering, Qatar University, Doha, Qatar

I. Ashraf

Department of electrical engineering, AMU, Aligarh 202002, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

95

S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,

Lecture Notes in Electrical Engineering 756,

https://doi.org/10.1007/978-981-16-0749-3_7

of these topologies have been investigated. However, the power quality and switching losses were the most concerned in such developments. Multilevel inverter are getting popularity in recent times over two-level inverters [3]. This is due to the advantages of high power processing capability of multilevel inverter with smaller device ratings along with better quality waveform at the output. In literature, several multilevel topologies have been investigated, and their performance have been compared on the basis of number of devices used as well as the number of devices used in the topologies. The main multilevel topologies that extensively investigated include neutral point clamped inverter (NPC), flying capacitor (FC), and cascaded H bridge inverter (CHB) [5]. Several new topologies are proposed in recent times, and the most promising are packed U-cell inverter, active neutral point clamped inverter (ANPC), modular multilevel converter (MMC), and quasi Z-source based multilevel inverter. Many new topologies with reduce device count have been reported in literature in recent years.

Several pulse width modulation (PWM) techniques for multilevel inverters have been proposed in literature for getting desired waveform with minimal undesired low-order harmonics component [3]. However, most of the proposed techniques were high switching frequency-based techniques such as level-shifted and phase-shifted carrier-based techniques, nearest level control and space vector PWM, etc. In high-power applications, high switching frequency results in high power losses, and thus, it is the main constraint in the high-power application of power converters. The preprogrammed pulse width modulation techniques such as selective harmonics elimination and selective harmonics minimization provide better quality output waveform with the minimum magnitude of selected low-order harmonics components [4]. For more levels in the output, the preprogrammed PWM techniques can have modulation and control at fundamental switching frequency and output waveform very close to the desired fundamental components of sinusoidal waveform. The output waveform is first synthesized, and then, Fourier series is used to derive expressions for fundamental component magnitude and harmonics component magnitudes [5, 6]. The system of nonlinear equations thus obtained have to be solved simultaneously for getting desired magnitude of the fundamental component and selectively remove the harmonics components.

For solving system of highly nonlinear and transcendental selective harmonics' minimization equations, different methods have been proposed in the literature [7]. The main methods are broadly categorized as numerical technique-based iterative methods, algebraic methods, and optimization-based evolutionary and meta-heuristic methods. The iterative techniques are highly dependent on initial guess, and algorithm may diverge if proper initial guess is not chosen [8]. Moreover, as the number of switching increased, the computation of derivative matrix (Jacobian matrix) becomes difficult, and singularity problems happened in subsequent iterations [9]. The algebraic methods are capable of producing exact and all the solutions but again as the switching angles increased, the order of polynomial also increased and it become very difficult to solve. In this paper, a continuous mode ant colony optimization (CMACO) technique is proposed to completely remove the certain low-order harmonics from the output. The computational results have been verified with the simulation and experimental results.

2 Problem Formulation for Harmonics Minimization

One leg of a generalized cascaded H-bridge inverter and its corresponding output stepped waveform is shown in Fig. 1. For having multiphase configuration, similar legs can be added, and the output waveform will shift by $\frac{2\pi}{P}$, where P is the number of phases. The Fourier analysis of quarterwave odd symmetric stepped waveform will result in mathematical output expression given as in (1).

$$v_{an}(\omega t) = \sum_{n=1}^{\infty} \frac{4V_{dc}}{n\pi} \left[\sum_{k=1}^S \cos(n\alpha_k) \right] \sin(n\omega t) \tag{1}$$

In the above expression, α_i is the switching instant of cell i , and S is the total number of individual cell and thus the total number of switching angles. Also n represents the harmonics order, and it is only odd valued since the even harmonics will be absent in the halfwave symmetric waveform. Also the P th order harmonics will be absent in the line voltage of P -phase configuration.

The switching angles are calculated only in a quarter period, and in other intervals, the switching angles are found accordingly as given in Table 1. If we have five H-cells with N_S separate DC sources, then any $N_S - 1$ harmonics can be controlled along with control of fundamental component. Let us consider three-phase H-bridge

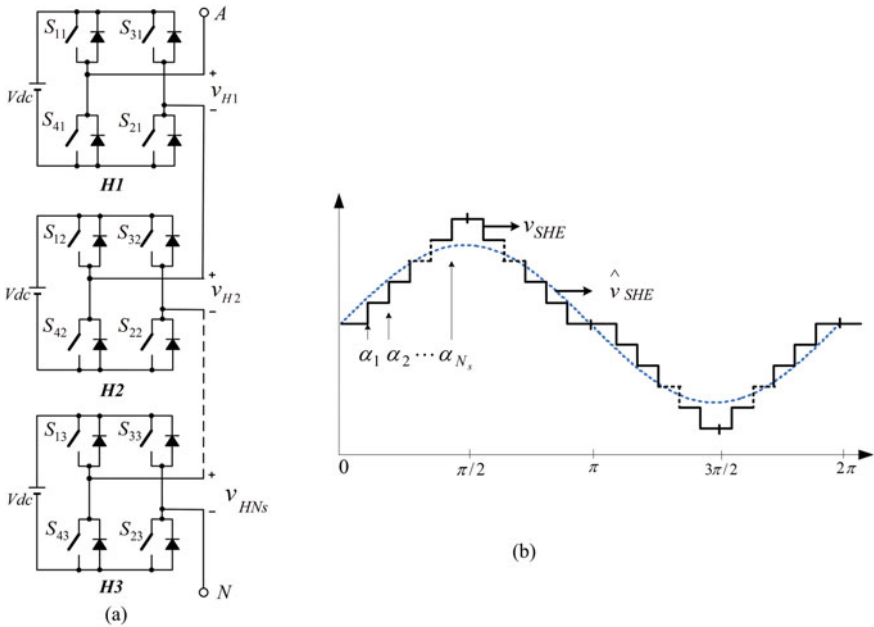


Fig. 1 a Generalized one leg of CHB circuit b Output stepped waveform

Table 1 Distribution of switching angles in quadrants

Quadrant	Switching angles
1st $\rightarrow (0, \frac{\pi}{2})$	$\alpha_1, \alpha_2, \dots, \alpha_{k-1}, \alpha_k$
2nd $\rightarrow (\frac{\pi}{2}, \pi)$	$\pi - \alpha_k, \pi - \alpha_{k-1}, \dots, \pi - \alpha_2, \pi - \alpha_1$
3rd $\rightarrow (\pi, \frac{3\pi}{2})$	$\pi + \alpha_1, \pi + \alpha_2, \dots, \pi + \alpha_{k-1}, \pi + \alpha_k$
4th $\rightarrow (\frac{3\pi}{2}, 2\pi)$	$2\pi - \alpha_k, 2\pi - \alpha_{k-1}, \dots, 2\pi - \alpha_2, 2\pi - \alpha_1$

inverter with five separate DC supplies. The harmonics elimination equations are have been established by considering elimination of 5th, 7th, 11th, and 13th order harmonics from the output and control of fundamental component by varying the modulation index, m .

3 Ant Colony Optimization (ACO)

The cooperative behavior of ants to find the shortest path in search of food and to form its colony using pheromone laying mechanism is called ant colony optimization. It is a meta-heuristic approach and is very useful in solving many optimization problems specially discrete optimization problems. The 'pheromone' information from the artificial ants is used to update toward better solution in successive iterations. The various steps involved are parameters initialization, random solutions but guided by pheromone information, optional local search, pheromone update, and finally check for termination criterion.

3.1 Solving Method by ACO

The ants which collect information during optimal path in search of food is stored in the form of pheromone trails represented with γ . Various terminologies of evolutionary algorithms are used by ACO as well. At the end of each generation, all the ants complete its tour, and the pheromone trail details are updated. Different variants of ACO algorithm are reported in literature based on the various pheromone trail update rules. For instance, the probability of ant i moving from node l toward node m in generation k is given by (2).

$$P_{l,m}^i(k) = \frac{\gamma_{l,m}(k) d_{l,m}^{-\mu}}{\sum_{x \in \psi_l^i} \gamma_{l,x} d_{l,x}^{-\mu}}, \quad m \in \psi_l^i \quad (2)$$

In, $\gamma_{l,m}$ is pheromone intensity on the edge $l \rightarrow m$, $d_{l,m}$ is distance between nodes l and m , ψ_l^i is all set of nodes that remain to be visited by, and i currently positioned at

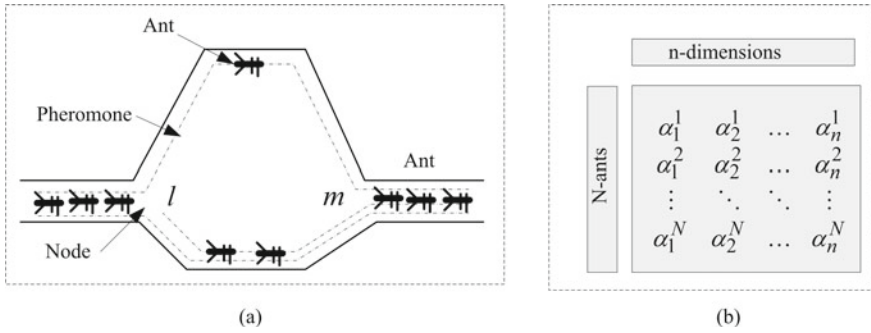


Fig. 2 a Decision making by ants b Ants solution space

node l and $\mu > 0$. An information storage list namely tabu list is used to keep records of the ants which already visited the particular node during each generation, and it is updated when the tour is completed. Once all the ants competed their tour, the pheromone information of all edges between node $l \rightarrow m$ is updated as per, where $\gamma_{l,m}^i(k)$ is the pheromone intensity on the path between $l \rightarrow m$ created by ant i and ς is pheromone decay parameter, $\varsigma \in (0, 1)$.

$$\gamma_{l,m}(k + 1) = (1 - \varsigma)\gamma_{l,m}(k) + \sum_{i=1}^N \gamma_{l,m}^i(k) \quad (3)$$

The pheromone intensity on $l \rightarrow m$ is taken to be $\frac{Q}{L_i}$ if ant i passes through this path, otherwise it is 0 as defined in (3), here Q is a constant. N is the number of ants defined in the algorithm. It results in a shorter route with stronger pheromone intensity. The pheromone evaporation phenomena is used to avoid local convergence of the ACO algorithm and therefore enables it to search for wider range of potential solutions. At last, a pheromone renewal is implemented using (4).

$$\gamma_{l,m}(k + 1) \leftarrow \max [\gamma_{\min}, \gamma_{l,m}(k + 1)] \quad \forall (l, m) \quad (4)$$

The decision making of ants between the nodes based on pheromone information is shown in Fig. 2a, whereas the search of the ants in the whole search space is shown in Fig. 2b. A pseudocode for writing programs in solving SHE problem is given in Table 2.

3.2 ACO Computational Results

Computation of switching angles with ant colony optimization technique is to formulate an objective function, and then, it is minimized. The individual ants contain

Table 2 ACO pseudocode

```

Define objective function for minimization  $f(\alpha)$ ,  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$ .
Initialize ACO parameters,  $r=1$ , and ant numbers (N).
    while (criterion)
        for ( $t = 1$  to  $N$ )
            From current node  $l$  select next node  $m$  defined by
            Check if all the path have been constructed
            end for
            update the pheromone information laid in
            Update the iteration,  $r = r + 1$ 
        end while
    Obtain the final results.

```

Table 3 ACO parameters

Parameters	Values
Maximum number of iterations	700
Population size (archive size)	10
Sample size	100
Intensification factor	0.5
Deviation-distance ratio	1

the potential solutions. The objective function is derived here by considering more weighting to regulate fundamental component and is given by. The ACO algorithm parameter used to solve the problem is given in Table 3.

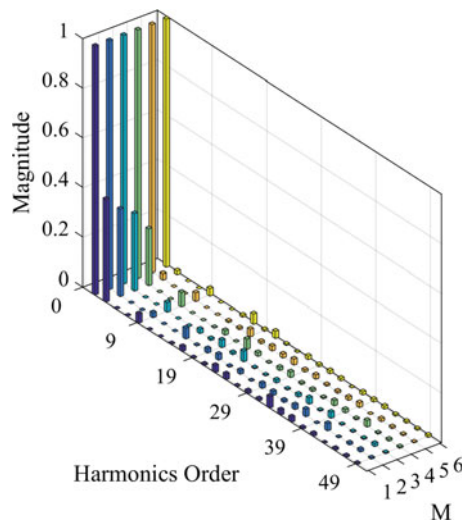
$$F = (10 \times (V_1^* - V_1))^4 + V_5^2 + V_7^2 + \dots + V_{2N-1}^2 \text{ or } V_{3N-2}^2 \quad (5)$$

Various cases have been computed using ant colony optimization to reduce certain low-order harmonics minimization from the output voltage. The ACO algorithms are mostly used for discrete optimization; however, here modified continuous ACO algorithm is implemented to compute the switching angles. The convergence rate is not as good as in case of GA and PSO algorithms. Selected computational results of switching angles, value of objective functions, and total harmonics distortion are given in the table. The harmonics profiles of these solution are shown in Fig. 3. From the results, it can be seen that the objective functions value is very close to zero, and the targeted harmonics are completely removed from the output. This is the continuous ACO which has advantage over discrete ACO in solving the selective harmonics minimization problem. Moreover, the implementation of continuous ACO is much simpler than discrete ACO method.

Table 4 Switching angles

M	α_1	α_2	α_3	α_4	α_5	Error	%THD
0.55	0.587	0.786	0.932	1.138	1.344	8×10^{-28}	7.09
0.60	0.459	0.761	0.902	1.085	1.263	2×10^{-31}	6.83
0.62	0.401	0.698	0.919	1.042	1.243	4×10^{-29}	7.21
0.67	0.306	0.564	0.868	1.009	1.190	2×10^{-30}	5.82
0.79	0.135	0.337	0.513	0.827	1.102	1×10^{-24}	5.68
0.81	0.095	0.327	0.434	0.744	1.065	5×10^{-28}	5.76

Fig. 3 Phase voltage harmonics profile



4 Simulation and Hardware Results

The simulation model for cascaded H-bridge converter have been developed using MATLAB/SIMULINK to verify the computational results. The FFT analysis is used to obtain the harmonics profile (Fig. 4).

Switching angles are first converted into time equivalent using $\alpha = \frac{360}{2\pi} \times T$, T being the time period. The percentage pulse width of pulses applied to the switch is then calculated using $\frac{2\pi}{(\alpha_{i+1}-\alpha_i)} \times 100$. Various cases for different switching angles have been tested, and simulation results found in very close agreement with the computational results. The pole voltage, line voltage, and harmonics spectrum of phase voltage for $M = 0.67$ with switching angles, $\alpha_1 = 0.038, \alpha_2 = 0.221, \alpha_3 = 0.367, \alpha_4 = 0.638, \alpha_5 = 0.696$, are given. The targeted harmonics for removal from the output waveform is absent in harmonics profile. The triplen harmonics will not appear in the line voltage harmonics profile for a balanced load as it will be automatically canceled.

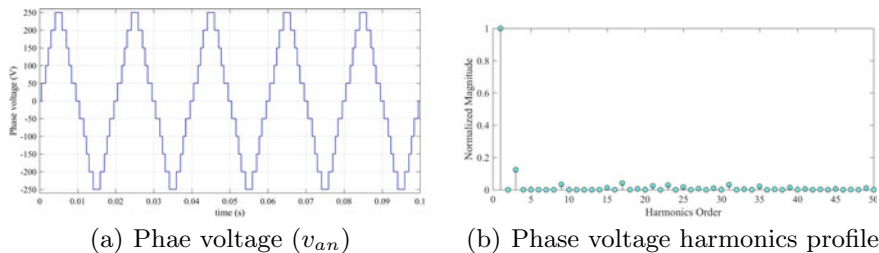


Fig. 4 11-level cascaded H-bridge waveforms $M = 0.69$, (2nd set)

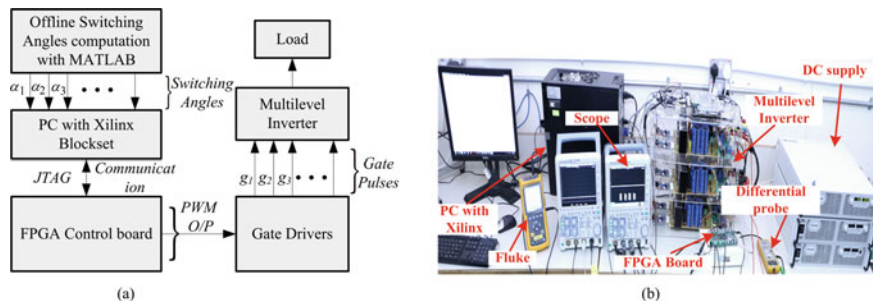


Fig. 5 Hardware setup for experimental results

The schematic and actual hardware setup of developed prototype in the laboratory is shown in Fig. 5. In this setup, separate programmable DC supplies are used to feed individual H-cells to generate 11-level stepped output waveform at fundamental switching frequency. The VHDL code for pulse width modulated waveform is generated from the PC and is applied to the gate drivers of the IGBT switches used to construct the H-bridge. The hardware result for the 11-level stepped voltage waveform is shown in Fig. 6 which is obtained from the prototype developed in the laboratory. An FPGA controller FPGA (VIRTEX-5 XC5VLX50T) is used to generate gate pulses for the IGBT semikron switch (SKM100GB12T4) based multilevel inverter. A fluke is used in conjunction with the differential probe to measure the harmonics profile of the output voltage waveform. The targeted harmonics for elimination such as 5th, 7th, 11th, and 13th are absent in the harmonics profile. Following the similar approach, the hardware result of any level and number of phases can be obtained.

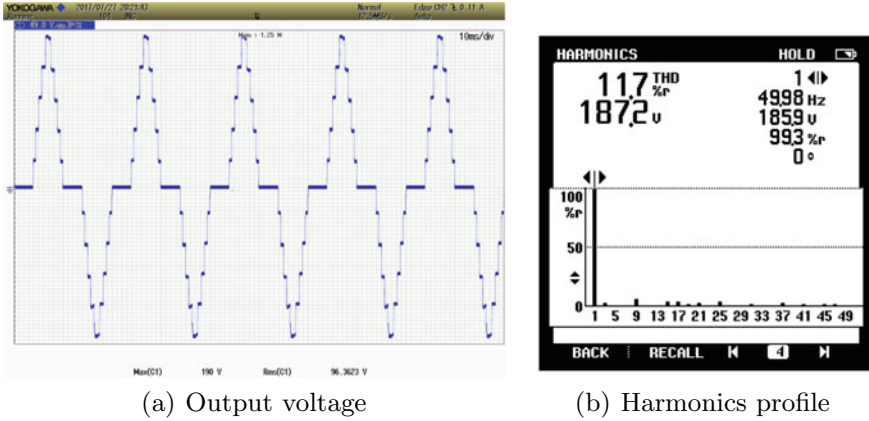


Fig. 6 11- level output voltage and harmonics profile

5 Conclusion

In this paper, a novel meta-heuristic technique based ant colony optimization in continuous mode implementation to minimize the selective low-order harmonics from the output of multilevel inverter is proposed. The computational results show the superiority of the proposed method over the discrete mode ant colony optimization method. The targeted harmonics are completely removed, and objective function reaches up to the order of 10^{-31} . Selected computational results for 11-level cascaded H-bridge inverter is given to illustrate the method. Computational, simulation, and hardware results confirm the practical applicability of the computed switching angles and elimination of targeted harmonics from the output.

References

1. M.H. Rashid, *Power Electronics Handbook* (Butterworth-Heinemann, 2017)
2. S. Kouro, M. Malinowski, K. Gopakumar, J. Pou, L.G. Franquelo, B. Wu, J. Rodriguez, M.A. Perez, J.I. Leon, Recent advances and industrial applications of multilevel converters. *IEEE Trans. Ind. Electron.* **57**(8), 2553–2580 (2010)
3. D.G. Holmes, T.A. Lipo, *Pulse Width Modulation for Power Converters: Principles and Practice* (Wiley, 2003)
4. S. Barkat, E.M. Berkouk, M.S. Boucherit, Particle swarm optimization for harmonic elimination in multilevel inverters. *Electr. Eng.* **91**, 221 (2009)
5. S. Ahmad, M. Meraj, A. Iqbal, M.A Alhitmi, Selective harmonics elimination in multilevel inverter by a derivative free iterative method under varying voltage condition. *ISA Trans.* Elsevier **92**, 241–256
6. S. Ahmad, Al-Hitmi M, A. Iqbal, I. Ashraf, M. Meraj, S. Padmanaban, Low-order harmonics control in staircase waveform useful in high power application by a novel technique. *Int. Trans. Electr. Energy Syst.* Wiley **29**(3) (2018)

7. S. Ahmad, Z.A. Ganie, I. Ashraf, A. Iqbal, Harmonics minimization in 3-level waveform and its FPGA realization. IEEE CIPECH, KIET, India. November 1-2 (2018). <https://doi.org/10.1109/CIPECH.2018.8724136>
8. S. Ahmad, A. Iqbal, R. Al-Ammari, H. Abu-Rub, Selected harmonics elimination in multi-level inverter using improved numerical technique, in *IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG 2018)*, pp. 1–6
9. S. Ahmad, I. Ashraf, A. Iqbal, M. Fatmi, SHE PWM for multilevel inverter using modified NR and pattern generation for wide range of solutions, in *IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG 2018)*, pp. 1–6

Loss Allocation Method for Microgrids Having Variable Generation



Dibya Bharti 

Abstract With some assumptions and limitations, various methods have been developed in literature mainly for loss allocation in transmission network and afterwards extended for radial distribution network, and some methods are specifically developed for radial distribution network. But, these methods are not suitable for microgrids which are integrated with conventional grid at sub-transmission and distribution levels depending on their geographical location. This paper presents a loss allocation method based on power flow results and relative position of buses for interconnected microgrid which is very effective in case of frequent change of generations due to intermittent nature of renewable resources. The implementation of the proposed method is very simple in microgrid with both meshed as well as radial topology without any computational complexity and requires only power flow solution and network data. The results are illustrated for different generating conditions of renewable sources in microgrid to demonstrate the efficiency of proposed loss allocation method.

Keywords Loss allocation (LA) · Microgrid · Variable generation · Power flow solution · Relative position

1 Introduction

Microgrid is an integration of various distributed generation (DG) especially renewable energy sources such as photovoltaic and wind which operate autonomously or in synchronous with conventional electrical grid. The incorporation of microgrids into distribution system has also transformed the structure of grid from radial to weakly meshed network. Energy insecurity, climate change and pollution are major concerns addressing significant changes in energy infrastructure by integrating renewable energy generation. In modern power system structure, several renewable generations are integrated to conventional grid at sub-transmission level and several may

D. Bharti (✉)

Department of Electrical and Electronics Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, UP 201009, India

be connected at distribution level. Hydro plant and wind farm are always far away from populated area and need to be connected to long-distance transmission. Some of solar generations are present at low-voltage distribution level. Due to independent ownership of DGs present in microgrid, it is essential to have a robust loss allocation (LA) method for attaining transparency. LA method should be applicable to both radial and meshed structured microgrid because depending on type of integration, some of the microgrids are of radial topology and some are of meshed topology.

1.1 Literature Review

In microgrids, conventional power generations are required to avoid power interruption as electricity generation by renewable energy resources is intermittent [1]; the output from a wind farm or a photovoltaic array depends on the climatic conditions. In interconnected mode, microgrid is connected with distribution network and works in co-ordination with the distribution management system. The power flow pattern varies very frequently in microgrids due to integration of non-conventional generation units. The presence of multiple source changes the distance between sources and loads which also alters network usages. Any LA method intended to be used for microgrid operation must be equally applicable to both radial as well as meshed network topology since the microgrid can be of either configuration. This makes LA problem very significant in microgrids. A robust LA method is required to differentiate between the contributions of individual participants, i.e. the generators and loads connected in the microgrid, and it should also consider amount of network usage of any participant as well as distance from source while making allocations to it.

There exist various LA methods in literature for transmission systems [2]. With some assumptions and modifications, transmission LA methods can also be used for distribution systems. The following part of this section discusses different allocation methods proposed earlier.

Pro-rata method distributes 50% loss to generator and 50% to loads, and then allocation is characterized by electric loss proportionally to the power delivered by each generator and each load but neglects their relative location within the network [2]. The inadequacy of pro-rata method was overcome by MW-mile method [3, 4], which considers the relative position of each participant from slack bus. Both pro-rata- and distance-based MW-mile methods are easy to understand and implement; however, these methods neglect the amount of power flowing through the network. If amount of power flowing through the branches changes, keeping the total power demand same, these methods will allocate similar loss to the different participants. In incremental/marginal transmission loss (ITL) allocation method, incremental loss coefficients are derived using Jacobians of NR method [5, 6]. ITL methods are suitable for networks with high X/R ratio and dependent on choice of slack bus [7]. Due to dependency on slack bus, ITL method results in over-recovery of network loss. But, in microgrid, usually conventional grid is connected at slack bus. Large amount

of loss will be allocated to slack bus that will be reallocated to other nodes present in conventional grid which is not reasonable. Based on results of power flow by NR method, direct loss coefficient (DLC) method allocates loss directly by establishing relation between real/reactive power of a bus and network loss [8]. Furthermore, applications of Hessian and Jacobian matrices are included in procedure of DLC and ITL methods, respectively, whose handling for larger system is computationally exhaustive. Proportional sharing principle-based methods [9, 10] allocate total network loss to either generators/DGs or loads as it involves application of linear principle.

A LA method dependent on Z-bus matrix of network is associated with non-singularity of admittance matrix of network and is not applicable to distribution networks where Y-bus matrix becomes singular [11]. Z-bus matrix method is not applicable to microgrid with radial topology when shunt admittance of the lines is negligible. LA by every method contains some degree of unreliability which heightens to explore economically and technically justified strategy for LA [11, 12]. Some methods integrate concepts of circuit theories with network characteristics to allocate loss in transmission system [13–20]. These methods are suitable for system with high value of X/R ratio and are not justified for LA in microgrids connected at distribution level where X/R ratio is low.

Some LA methods are specially proposed for distribution systems [21–28]. Loss caused by each branch current is decomposed into nodal injection which requires an additional forward sweep power flow on modified network consisting only branch resistances in [21]. Carpeno et al. characterize the different LA techniques for radial distribution configuration with DGs [22]. Method proposed by Costa requires running power flow twice for allocating loss to loads and DGs in first two steps, and remaining loss is redistributed to DGs only in proportion to their apparent power [23]. Atanasovski proposed tracing-based approaches for allocating loss in radial distribution system based on disintegration of power [24] and energy [25] into respective nodal injections. For radial distribution system, Savier also developed methods for allocating branch loss to the nodes by using exact formulation [26, 27]. Branch-oriented procedure is proposed by Jahromi in which loss is allocated to loads and DGs separately, and normalization is applied to avoid over-recovery of loss [28].

LA approaches adapted by Jagtap are either current summation or power summation and branch oriented for radial distribution network [29–31]. These methods require only power flow solution to allocate loss to DGs and loads and employ backward sweep network reduction technique to remove the steps of normalization.

Most of the proposed methods [21–36] are particularly suggested for radial systems and are applicable only for microgrid with radial topology. A game theory-based LA method is proposed for radial and weakly meshed networks which overcomes the limitations of conventional Shapley value method but considers DGs as negative loads [37–40].

1.2 Contribution of Proposed Work

A fair and satisfactory LA method should reflect both the network topology and the magnitude of power injected or consumed at a bus. The present work proposes a LA strategy for interconnected microgrid with meshed as well as radial structure which works on power flow result of the system and relative distance between buses in network. Proposed method requires only power flow results and electrical closeness centrality indices which is a measure of the degree to which an individual is near all other individuals in a network. The proposed LA method is straightforward without any intricate computational applications.

1.3 Organization of Paper

The next section confers electrical closeness centrality measures and their use in proposed LA method. Section 2 discusses relevance of electrical closeness centrality measure in LA in brief and describes the steps of calculating closeness centrality and proposed method with an example. Section 3 demonstrates the application of proposed method in different scenario of microgrid. Finally, Sect. 4 concludes the work.

2 Centrality Measures and Loss Allocation

Distributed energy generations (DERs) of microgrids are owned by different entity which necessitates implementation of a robust LA method by distribution system operator (DSO). In microgrids, generations are sporadic subject to climatic changes which lead to change in power flow results. With changing generation pattern, the power flows through different paths of the network changes, resulting into change in network usage. Electrical closeness centrality is the measure of relative position of a bus in the network and dependent on system data and power flow results. To change the loss contribution according to relative position and network usage by individual participant, electrical closeness centrality is used for allocating loss to generators and loads.

Electrical closeness centrality measures are calculated by using bus dependency matrix [41–43], which exhibits dependability of buses on each other present in network. The method discussed for finding bus dependency matrix in [39] is applicable to both meshed and radial electrical network. A fair LA method needs to take care of the relative location and importance of any bus in the network, and this aspect can be included by incorporating centrality index in the allocation method. The following section discusses the method of determining centrality measure of a network.

2.1 Calculation of Electrical Closeness Centrality Measures

The electrical closeness centrality measure is calculated from the bus dependency matrix which is of order (bus * bus). The bus dependency matrix of any n -bus system is calculated as Eq. (1) then,

$$D_{\text{bus_dep}_n} = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \dots & \dots & \vdots \\ d_{n1} & d_{n2} & \dots & d_{nn} \end{bmatrix} \quad (1)$$

Electrical closeness centrality of each can be calculated by row sum of matrix. For example, closeness centrality of n th bus will be calculated as

$$C_n = d_{n1} + d_{n2} + \dots + d_{nn} \quad (2)$$

The closeness centrality computes the extent of connectivity to which a bus is in close proximity to all other buses of the network. Electrical closeness centrality measures are calculated by using bus dependency matrix which is dependent on system impedance and power flow results. A fair LA method should incorporate the relative location and importance of any bus in the network, and this aspect can be included by integrating centrality index in the allocation method.

2.2 Loss Allocation Using Electrical Closeness Centrality Measures

The algorithm used for allocating loss to each nodes of the network is given as follows:

- Step I: Calculate total loss (P_{loss}) of the system using power flow results.
- Step II: Calculate electrical closeness centrality indices (C_i) for each bus present in the network by using bus dependency matrix as discussed in Sect. 2.1.
- Step III: Calculate total outgoing power (P_i) at each bus present in the system.
- Step IV: Calculate proportional indices (a_i) for each bus by using Eq. (3). where ' N ' is the number of buses present in network.

$$a_i = \frac{P_i}{\sum_{i=1}^N P_i C_i} \quad (3)$$

Step V: Loss allocated to each bus can be given by

$$p_i = P_{\text{loss}} \cdot a_i \cdot C_i \quad (4)$$

Step VI: Normalize loss at various buses to calculate loss contribution of individual loads and generators.

2.3 Validation of Proposed Loss Allocation Method

A test system with five-bus and seven links is considered as microgrid with meshed topology to demonstrate the applicability of proposed method. Microgrid is integrated with conventional grid at bus 1, and there are two non-conventional energy resources: solar plant and wind farm with installed capacity of 40 MW and 30 MW, respectively, connected at bus 2 and bus 3. Figure 1 represents five-bus test system with system impedance, outflow power and inflow power at each node. The bus dependency matrix (given below in Eq. (5)) is calculated which depends on the shortest path and power flow result. Total loss of the system (P_{loss}) is 12.6806 MW.

$$D_{bus_dep} = \begin{bmatrix} 0.0000 & 4.8650 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.9803 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 10.7646 & 0.0000 \\ 0.0000 & 1.9730 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 4.9730 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix} \quad (5)$$

As described above, the row sum of bus dependency matrix can be used as electrical closeness centrality measures which are $C_1 = 4.8650$, $C_2 = 0.9803$, $C_3 = 10.7646$, $C_4 = 1.9730$ and $C_5 = 4.9730$ and outgoing power are $P_1 = 142.6805$, $P_2 = 149.3261$, $P_3 = 64.7077$, $P_4 = 111.1685$ and $P_5 = 60.0000$ for buses 1, 2, 3, 4 and 5, respectively.

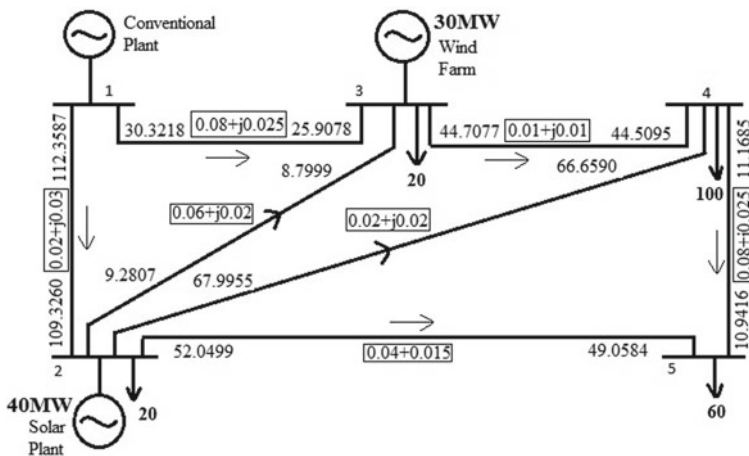


Fig. 1 Meshed structured microgrid with five buses and seven links

By using Eq. (4), loss allocated to buses 1, 2, 3, 4 and 5 can be calculated as $p_1 = 4.2837$, $p_2 = 0.9034$, $p_3 = 4.2985$, $p_4 = 1.3536$ and $p_5 = 1.8414$, respectively. The fairness of the LA method can be verified as $p_1 + p_2 + p_3 + p_4 + p_5 = 12.6806$ MW.

The proposed LA technique can be easily applied to the microgrids where change in generations is very frequent as it requires very less computational time and is a very fast and efficient technique. Electrical closeness centrality measures-based LA method is efficiently applicable to both meshed structured microgrid and radial topology-based microgrid.

3 Application of Proposed Method

To illustrate applicability of proposed method with variable sources, modified IEEE 14-bus test system is considered as an interconnected microgrid with meshed topology. Two different cases with altered renewable generation are considered to validate the applicability of proposed method in microgrids where alteration in power flow pattern is very frequent. The following scenarios are considered as microgrid for analysis of proposed LA method.

3.1 Application of Proposed Loss Allocation Method with Variable Sources

Case 1: Interconnected microgrid with one solar plant and one wind farm

A modified IEEE 14-bus system is considered in which node 2 is assumed to be connected with a wind farm of 40 MW rated capacity and node 3 has a concentrated solar plant of 60 MW rated capacity. A new branch is added between buses 1 and 3 in this modified system, and conventional grid is interconnected at bus 1. Single line diagram of the modified test system is shown in Fig. 2 with direction of power flow through the lines. Branch data for the modified system are listed in Table 1.

After power flow, total loss of the system (P_{loss}) is calculated as 4.4700 MW. After power flow, by using LA technique proposed in Sect. 2.2, loss allocated to each bus present in the network is given in Table 2.

In Table 2, electrical closeness centrality, outgoing power and loss allocated to each bus are listed. The method for calculating electrical closeness centrality by bus dependency matrix is briefly discussed in Appendix A. In Table 2, loss allocated to bus 8 is zero as neither generator nor load is connected at bus 8 as shown in Fig. 2. From Fig. 2, it can also be observed that generator connected at bus 6 is not delivering any power but load is obtaining power. So, the loss is being allocated to bus 6. From results presented in Table 2, it can be concluded that total losses allocated to generators and loads are 2.463 MW and 2.0068 MW, respectively. So, the proposed method is not dividing total loss to generators in equal proportional unlike pro-rata

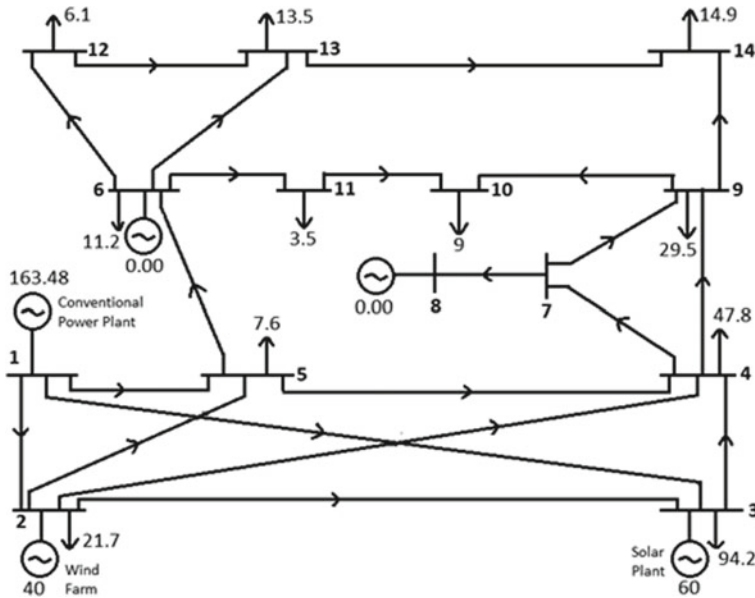


Fig. 2 Modified IEEE 14-bus test system as meshed structured microgrid

method, and this is realized by including electrical closeness centrality indices which is the representation of relative electrical distance of buses present in network.

Case II: Interconnected microgrid with two solar plants and two wind farms

Again, modified IEEE 14-bus system is considered as microgrid in which both node 2 and 6 are assumed to be connected with a wind farm of 40 MW rated capacity, and node 3 and 8 have a concentrated solar plant of 60 MW and 50 MW rated capacity, respectively. After power flow, total loss of the system (P_{loss}) is calculated as 2.8777 MW. Proposed LA method is applied to IEEE-14 bus system considered as microgrid, and results are listed in Table 3.

From the results of Table 3, it can be concluded that total loss allocated to generators is 1.6671 MW while total loss allocated to loads is 1.2106 MW which indicates that total loss is not being shared by generators and loads in equal fraction which was drawback of LA methods related to pro-rata techniques.

From Tables 2 and 3, it is clear that depending on power flow results, electrical closeness centrality changes. Loss allocated to buses changes depending on power flow results and electrical closeness centrality but sum of loss allocated to buses is exactly equal to the total loss of the system.

Table 1 Branch data of modified IEEE 14-bus system

From bus	To bus	Resistance (in Ω)	Reactance (in Ω)
1	2	0.01938	0.05917
1	5	0.05403	0.22304
1	3	0.00000	0.04211^a
2	3	0.04699	0.19797
2	4	0.05811	0.17632
2	5	0.05695	0.17388
3	4	0.06701	0.17103
4	5	0.01335	0.04211
4	7	0.00000	0.20912
4	9	0.00000	0.55618
5	6	0.00000	0.25202
6	11	0.09498	0.19890
6	12	0.12291	0.25581
6	13	0.06615	0.13027
7	8	0.00000	0.17165
7	9	0.00000	0.11001
9	10	0.03181	0.08450
9	14	0.12711	0.27038
10	11	0.08205	0.19207
12	13	0.22092	0.19988
13	14	0.17093	0.34802

^aNewly added line

3.2 Application of Proposed Loss Allocation Method in Microgrid with Radial Topology

A microgrid with radial topology is considered in Fig. 3. It represents a modified 12-bus radial system integrated with conventional power plant at bus 1, and a solar plant and wind farm are connected at bus 5 and 9, respectively. It is assumed to be connected with a wind farm of 60 MW rated capacity at bus 9, and node 5 has a concentrated solar plant of 90 MW rated capacity. Bus 1 is not connected to any load; only conventional power plant is integrated into the network at bus 1. All the connected loads (shown in Fig. 3) are in MW. Direction of power flow is shown in Fig. 3.

After power flow, it has been found that the total loss of the system is 6.8429 MW. The loss allocated by the proposed method is listed in Table 4.

Table 2 Results of LA of modified IEEE 14-bus system (with two DERs) considered as microgrid

Bus	Electrical closeness centrality	Outgoing power (in MW)	Loss allocated (in MW)
1	9.9971	163.4800	1.0149
2	9.8958	96.4100	0.5924
3	10.8397	127.1300	0.8557
4	9.0000	94.3700	0.5274
5	8.9733	73.5900	0.4101
6	8.9777	41.6400	0.2321
7	9.9524	29.6100	0.1830
8	11.9524	0.0000	0.0000
9	9.9465	46.5700	0.2876
10	10.9456	9.0000	0.0612
11	10.9963	5.8200	0.0397
12	10.9993	7.5300	0.0514
13	9.9993	18.2100	0.1131
14	10.9407	14.9000	0.1012
Total			4.4700

Table 3 Results of LA of modified IEEE 14-bus system (with four DERs) considered as microgrid

Bus	Electrical closeness centrality	Outgoing power (in MW)	Loss allocated (in MW)
1	10.8214	71.8779	0.3613
2	8.9686	86.0742	0.3586
3	10.9765	94.2000	0.4803
4	8.9801	58.9536	0.2459
5	8.9929	39.8391	0.1664
6	8.9440	45.3010	0.1882
7	9.9765	66.0341	0.3060
8	12.0000	50.0000	0.2787
9	9.9684	43.0276	0.1992
10	10.9715	9.0000	0.0459
11	10.9717	8.0126	0.0408
12	10.9748	7.7963	0.0397
13	9.9852	19.5840	0.0908
14	10.9666	14.9000	0.0759
Total			2.8777

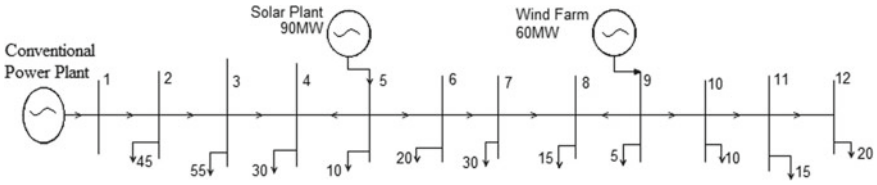


Fig. 3 Twelve-bus radial system modified as microgrid

Table 4 Results of loss allocation of 12-bus radial system considered as microgrid

Bus	Electrical closeness centrality	Outgoing power (in MW)	Loss allocated (in MW)
1	15.8263	111.8429	1.6618
2	10.8454	110.6288	1.1264
3	8.9293	64.7067	0.5424
4	8.9751	30.0000	0.2528
5	8.8759	90.0000	0.7500
6	8.9586	56.6397	0.4764
7	8.9201	36.3377	0.3043
8	8.9603	15.0000	0.1262
9	8.9071	60.0000	0.5017
10	8.9676	45.0362	0.3792
11	12.3343	35.0049	0.4054
12	16.8483	20.0000	0.3164
Total			6.8429

From the results of Table 4, also it can be concluded that proposed method overcomes the drawback of LA methods based on pro-rata by sharing 2.9135MW to generators and 3.9295MW to loads.

It has been also observed from Tables 2, 3 and 4 that LA by proposed method has not given negative loss contribution to any generators or loads.

4 Conclusions

This paper proposes a simple and robust method for LA in interconnected microgrid with meshed topology as well as radial topology where generations are considered to be variable. Due to various non-conventional energy sources in microgrid, power flow changes very frequently and presence of multiple sources modifies network usages. Proposed method easily determines loss allocated to each bus for every scenario. While allocating loss to different buses, the proposed method considers relative position of buses in the network and requires only power flow solution with network

data. Electrical closeness centrality measure is used to identify the relative location of buses present in the network and can be calculated by using power flow results. The results obtained by proposed method show that electrical closeness centrality changes according to power flow results and corresponding to that amount of loss allocated to different buses changes. LA by proposed method includes position of each load, and generators in the network and loss contribution to generators and loads depend on the amount of power produced or consumed by them. The proposed procedure is simple to understand, and its execution is undemanding because it does not require intricate computational application.

APPENDIX: Calculation of elements of bus dependency matrix

The computation of bus dependency matrix depends on shortest path between pair of buses and active power flowing it. In shortest path between pair of bus of power system network, there will be intermediate buses if selected buses are not adjacent buses. There may be either single bus or multiple intermediate buses in shortest path. So, the elements of bus dependency matrix can be calculated by following the steps written below:

- Step I: Run power flow of the system.
- Step II: Determine the shortest path for each bus pair by assigning impedance as weights.
- Step III: Find the maximum active power in each shortest path, P_{st} . (See Fig. 4 and Fig. 5).

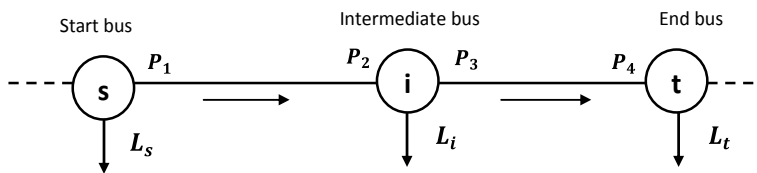


Fig. 4 Description of $P_{st(i)}$ and P_{st} in shortest path with single intermediate bus

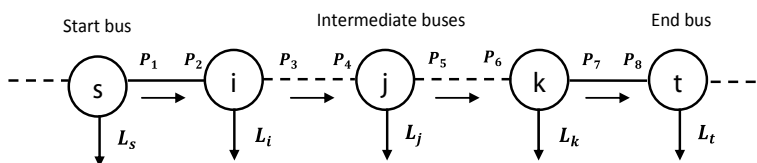


Fig. 5 Description of $P_{st(i)}$ and P_{st} in shortest path with multiple intermediate buses

Step IV: Ascertain the maximum of inflow and outflow at intermediate bus within each shortest electrical path, $P_{st}(i)$. (See Fig. 4 and Fig. 5).

Step V: The dependency of bus s upon bus i to transmit power to other buses of the network can be given by,

$$d_{si} = \sum_{t=1}^n \frac{P_{st}(i)}{P_{st}}$$

$$s \neq t \neq i \in v$$

$$P_{st} = \max of (P_1, P_2, P_3, P_4)$$

$$P_{st}(i) = \max of (P_2, P_3)$$

$$P_{st} = \max of (P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8)$$

$$P_{st}(x) = \max of [\max(P_2, P_3), \max(P_4, P_5), \max(P_6, P_7)]$$

x is the intermediate bus depending upon amount of active power inflow/outflow.

For example, in a 5-bus system, element d_{14} and d_{35} are calculated as:

$$d_{14} = \sum_{t \in \{2,3,5\}} \frac{P_{1t}(4)}{P_{1t}} = \frac{P_{12}(4)}{P_{12}} + \frac{P_{13}(4)}{P_{13}} + \frac{P_{15}(4)}{P_{15}}$$

$$d_{35} = \sum_{t \in \{1,2,4\}} \frac{P_{3t}(4)}{P_{3t}} = \frac{P_{31}(5)}{P_{31}} + \frac{P_{32}(5)}{P_{32}} + \frac{P_{34}(5)}{P_{34}}$$

The diagonal elements of bus dependency matrix will be zero and its row summation gives electrical closeness centrality.

References

1. P. Chakraborty, E. Baeyens, P. Khargonekar, Distributed control of flexible demand using proportional allocation mechanism in a smart grid: game theoretic interaction and price of anarchy. *Sustain. Energy Grids Netw.* **12**, 30–39 (2017)
2. A. Conejo, J. Arroyo, A. Guijarro, Transmission loss allocation: a comparison of different practical algorithms. *IEEE Trans. Power Syst.* **17**(3), 571–576 (2002)
3. D. Shirmohammadi, P.R. Gribik, Evaluation of network capacity use for wheeling transaction. *IEEE Trans. Power Syst.* **4**(4), 1405–1413 (1989)
4. H.H. Happ, Cost of wheeling methodologies. *IEEE Trans. Power Syst.* **9**(1), 147–156 (1994)

5. E. Carpaneto, G. Chicco, J.S. Akilimali, Computational aspects of the marginal loss allocation methods for distribution systems with distributed generation, in *Proceedings of IEEE MELECON*, Benelmadena (Malaga), Spain, pp. 1028–1031 (2006)
6. F.D. Galiana, A.J. Conejo, I. Kockar, Incremental transmission loss allocation under pool dispatch. *IEEE Trans. Power Syst.* **17**(1), 26–33 (2002)
7. F. Galiana, M. Phelan, Allocation of transmission losses to bilateral contracts in a competitive environment. *IEEE Trans. Power Syst.* **15**(1), 143–150 (2000)
8. J. Mutale, G. Strbac, S. Curcic, N. Jenkins, Allocation of losses in distribution systems with embedded generation. *Proc. IEEE Gener. Transm. Distrib.* **147**(1), 7–14 (2000)
9. G. Strbac, D. Kirschen, S. Ahmed, Allocating transmission system usage on the basis of traceable contributions of generators and loads to flows. *IEEE Trans. Power Syst.* **13**(1), 527–534 (1998)
10. J.W. Bialek, Tracing the flow of electricity. *Proc. IEEE Gener. Transm. Distrib.* **143**(4), 313–320 (1996)
11. A. Conejo, F. Galiana, I. Kockar, Z-bus loss allocation. *IEEE Trans. Power Syst.* **16**(1), 105–110 (2001)
12. A. Gómez, J.M. Riquelme, T. González, E.A.R. Velasco, Fair allocation of transmission power losses. *IEEE Trans. Power Syst.* **15**(1), 184–188 (2000)
13. S. Abdelkader, D. Morrow, A. Conejo, Network usage determination using a transformer analogy. *IET Gener. Transm. Distrib.* **8**(1), 81–90 (2014)
14. J. Daniel, R. Salgado, M. Irving, Transmission loss allocation through a modified Ybus. *IEE Proc., Gener. Transm. Distrib.* **152**(2), 208–214 (2005)
15. Q. Ding, A. Abur, Transmission loss allocation in a multiple transaction framework. *IEEE Trans. Power Syst.* **19**(1), 214–220 (2004)
16. S. Abdelkader, Characterization of transmission losses. *IEEE Trans. Power Syst.* **26**(1), 392–400 (2011)
17. Y. Molina, R. Prada, O. Saavedra, Complex losses allocation to generators and loads based on circuit theory and Aumann-Shapley method. *IEEE Trans. Power Syst.* **25**(4), 1928–1936 (2010)
18. K. Min, S. Ha, S. Lee et al., Transmission loss allocation algorithm using path-integral based on transaction strategy. *IEEE Trans. Power Syst.* **25**(1), 195–205 (2010)
19. S. Valerie, D. McDonald, K. Tapan, Development of a new loss allocation method for a hybrid electricity market using graph theory. *Electr. Power Syst. Res.* **79**(2), 301–310 (2009)
20. A. Elmitwally, A. Eladl, S.M. Abdelkader, Efficient algorithm for transmission system energy loss allocation considering multilateral contracts and load variation. *IET Gener. Transm. Distrib.* **9**(16), 2653–2663 (2015)
21. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON)* held at Deenbandhu Chhotu Ram University of Science and Technology, SONEPAT, India on Feb 28—March 1, 2020
22. E. Carpaneto, G. Chicco, J.S. Akilimali, Branch current decomposition method for loss allocation in radial distribution systems with distributed generation. *IEEE Trans. Power Syst.* **21**(3), 1170–1179 (2006)
23. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
24. E. Carpaneto, G. Chicco, J.S. Akilimali, Characterization of the loss allocation techniques for radial systems with distributed generation. *Electr. Power Syst. Res.* **78**(8), 1396–1406 (2008)
25. P. Costa, M. Matos, Loss allocation in distribution networks with embedded generation. *IEEE Trans. Power Syst.* **19**(1), 384–389 (2004)
26. M. Atanasovski, R. Taleski, Power summation method for loss allocation in radial distribution networks with DG. *IEEE Trans. Power Syst.* **26**(4), 2491–2499 (2011)
27. R.N. Shaw, P. Walde, A. Ghosh, A new model to enhance the power and performances of 4x4 PV arrays with puzzle shade dispersion. *Int. J. Innov. Technol. Explor. Eng.* **8**(12) (2019)

28. M. Atanasovski, R. Taleski, Energy summation method for loss allocation in radial distribution networks with DG. *IEEE Trans. Power Syst.* **27**(3), 1433–1440 (2012)
29. J. Savier, D. Das, An exact method for loss allocation in radial distribution systems. *Int. J. Electr. Power Energy Syst.* **36**(1), 100–106 (2012)
30. J. Savier, D. Das, Energy loss allocation in radial distribution systems: a comparison of practical algorithms. *IEEE Trans. Power Delivery* **24**(1), 260–267 (2009)
31. Z. Ghofrani-Jahromi, Z. Mahmoodzadeh, M. Ehsan, Distribution loss allocation for radial systems including DGs. *IEEE Trans. Power Delivery* **29**(1), 72–80 (2014)
32. K.M. Jagtap, D.K. Khatod, Loss allocation in radial distribution networks with different load models and distributed generations. *IET Gener. Transm. Distrib.* **9**(12), 1275–1291 (2015)
33. K.M. Jagtap, D.K. Khatod, Loss allocation in radial distribution networks with various distributed generations and load models. *Int. J. Electr. Power Energy Syst.* **75**, 173–186 (2016)
34. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol. 661, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
35. K.M. Jagtap, D.K. Khatod, Loss allocation in distribution network with distributed generations. *IET Gener. Transm. Distrib.* **9**(13), 1628–1641 (2015)
36. K.M. Jagtap, D.K. Khatod, Novel approach for loss allocation of distribution networks with DGs. *Electr. Power Syst. Res.* **143**, 303–311 (2017)
37. K.M. Jagtap, D.K. Khatod, Current summation based approach for loss allocation with distributed generation, in *Proceedings of the IEEE Power Instrumentation, Control and Computing (PICC)*, pp. 1–5 (2018)
38. P. Kumar, N. Gupta, K.R. Niazi, A. Swarnkar, A circuit theory-based loss allocation method for active distribution systems. *IEEE Trans. Smart Grid* **10**(1), 1005–1012 (2019)
39. P. Kumar, N. Gupta, K.R. Niazi, A. Swarnkar, Branch current decomposition method for loss allocation in contemporary distribution systems. *Int. J. Electr. Power Energy Syst.* **99**, 134–145 (2018)
40. H. Kumar, D.K. Khatod, Loss allocation for distribution network using τ -value game theory approach, in *Proceedings of the IEEE Intelligent Circuits and Systems*, pp. 401–405 (2018)
41. S. Sharma, A.R. Abhyankar, Loss allocation for weakly meshed distribution system using analytical formulation of Shapley value. *IEEE Trans. Power Syst.* **32**(2), 1369–1377 (2017)
42. S. Sharma, A.R. Abhyankar, Loss allocation of radial distribution system using Shapley value: a sequential approach. *Int. J. Electr. Power Energy Syst.* **88**, 33–41 (2017)
43. D. Bharti, M. De, A centrality index based approach for selection of optimal location of static reactive power compensator. *Electr. Power Compon. Syst. J.* **46**(8), 886–899 (2018)

Contingency Analysis of Power Network with STATCOM and SVC



P. S. Vaidya and V. K. Chandrakar

Abstract In the modern era, where the demand of electricity is increasing very rapidly, for urbanization of country power shortage is a main issue. This power shortage is becoming a nightmare issues for humankind. So, the interconnection of different grids is done. Interconnections of different sources may result in various stability issues due to fault occurrences or due to any outages of any power network elements. To stabilize the power network during such outage or contingency condition, power device elements are playing a crucial role. This paper shows the effect of using STATCOM and SVC as a power device for maintaining stability as well as voltage profile of power network. Power network simulations are done in PSAT, a toolbox of MATLAB for IEEE-9 bus network.

Keywords Power flow · Contingency · FACTS · STATCOM · SVC · IEEE 9 bus network · CPF · MATLAB/PSAT

1 Introduction

Today, the electricity is one of our fundamental needs. As there is an increase in the population day by day the demand of electricity also increases. Hence, the power network has to increase the generation capacity of existing network or to build new power network. The working of new power network is exceptionally costlier choice, so the upgradation of existing network is an elective alternative. The power network works close to its basic working points of confinement to fulfill the expanding load demand. With the increase in demand, power devices are another option to balance power and demands during various contingency. Control of power network is the most noteworthy problem to be concerned with. Day by day the network is getting

P. S. Vaidya (✉) · V. K. Chandrakar
Department of Electrical Engineering, G H Raisoni College of Engineering, Nagpur, India
e-mail: prajakta.vaidya@raisoni.net

V. K. Chandrakar
e-mail: vinod.chandrakar@raisoni.net

complex, and stability issues are getting to be vulnerable. Voltage instability issues by and large increases as the different stochastic energy sources like solar, wind and so on are joined to the network. For the secure operation limits of violations of the power network is monitored by network operator. We can state that the transmission clog happens, when the breaking points are violated because of sudden increment in load demand, generation outage or transmission line blackout and so forth. The network stability and security are influenced because of blockage. Voltage instability issues can be understood by giving satisfactory reactive power support at proper area in the network [1]. For these reasons, different compensating devices are utilized by utilities, every one of which has its own particular limitations and characteristics. Voltage stability is the capacity of the power network to keep up sufficient voltage size with the end goal that the actual power exchanged by network load to that will increment. PV curve is broadly utilized for examining steady and insecure condition. MATLAB toolbox, power system analysis is used for control and study of power network. It does analysis like power flow, stability study, time domain simulations and continuation power flow [2]. It is used for small and medium of static and dynamic analysis of power network in MATLAB [3].

2 Overview of STATCOM and SVC

Synchronous voltage source with least and most extreme voltage extent limits is implied as STATCOM and SVC. It is as shown in Fig. 1a, b.

To compensate the recipient power line voltage, replacing of shunt capacitor bank is used. STATCOM offers various favorable circumstances over banks of shunt capacitors, for example, considerably more tightly control of the voltage compensation and increased line stability during load variations [4–7]. Static VAR compensator is also a shunt FACTS device used to control the power flow in a power network. The static VAR compensator is made of the capacitor banks and air co-reactors connected in

Fig. 1 a STATCOM. b SVC

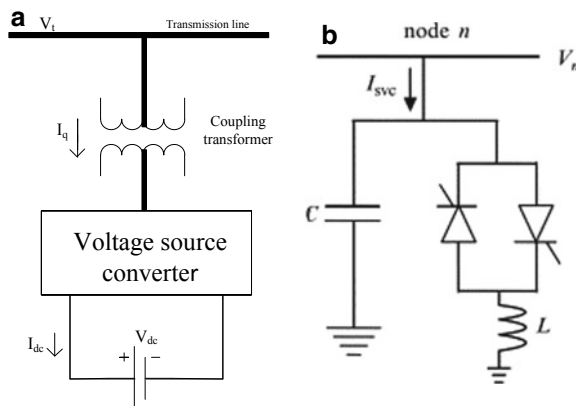
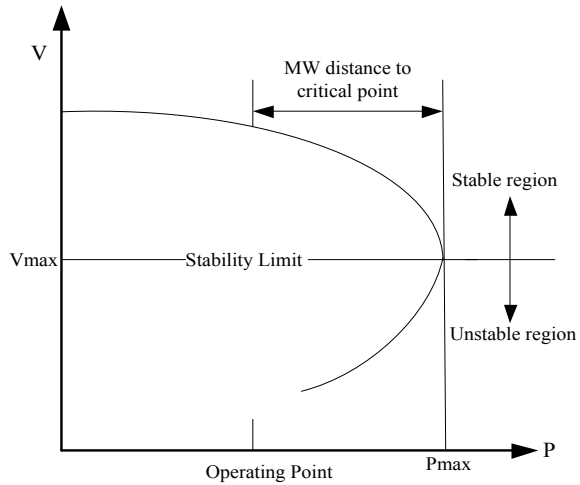


Fig. 2 Real power voltage (P–V) curve



shunt. The air core reactors are connected in series to thyristor. For all intents and purposes, a STATCOM or SVC is mounted to aid power networks that have a poor power factor and poor voltage control [8–10].

3 Contingency Analysis

Unpredictable condition in the power network is known as contingency. To maintain power network to be secure, it must have continuous power flow with no losses. Hence, the contingency analysis is performed to ensure the avoidance of crisis condition and to perform the network at low cost [11]. Whenever the pre determined breaking points of the power network gets out of limit the network is said to be in unbalanced condition. These violations of the breaking points result from unpredictable conditions in the network. Contingency in a power network results in instability of complete power network and affects the reliability, security and continuity of network. Outage of a transformer, generator and or line, is termed as contingency [12–17].

The N-R tool is well accepted due to its minimum iteration with rapid convergence. The N-R method is flexible; hence, the load flow forms can optimize power network operation.

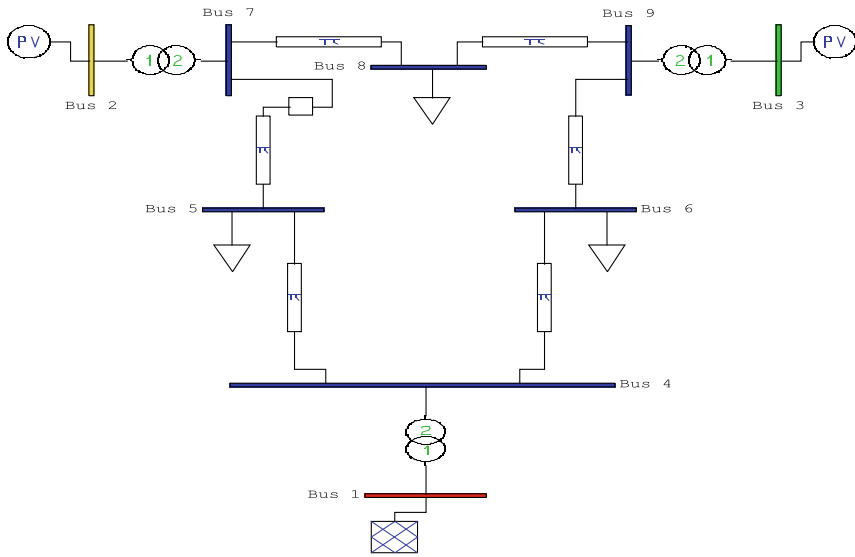


Fig. 3 9 bus IEEE network without FACTS device

4 PV Curves Analysis

Figure 2 shows the PV curve, the transform in power from one bus to another bus which influences the bus voltages are given by the PV curve [18–21]. The power network is widely operated in the upper part of PV curve.

5 Test Network

Network consists of one slack bus, two PV generators connected in Bus 1, Bus 2 and Bus 3, respectively, as shown in Figs. 3, 4 and 5. The network model consists of three step-up transformers connected between Bus 1–4, Bus 2–7 and Bus 3–9. The three loads are connected in buses 5, 6 and 8. The power device, i.e., STATCOM or SVC, is connected in Bus 5. MATLAB/PSAT software is used for simulation.

The details of test network components are as in Tables 1 and 2.

6 Results

Each line is removed one by one like as a contingency condition in PSAT software, which results in poor performance of power network. Continuous power flow is done

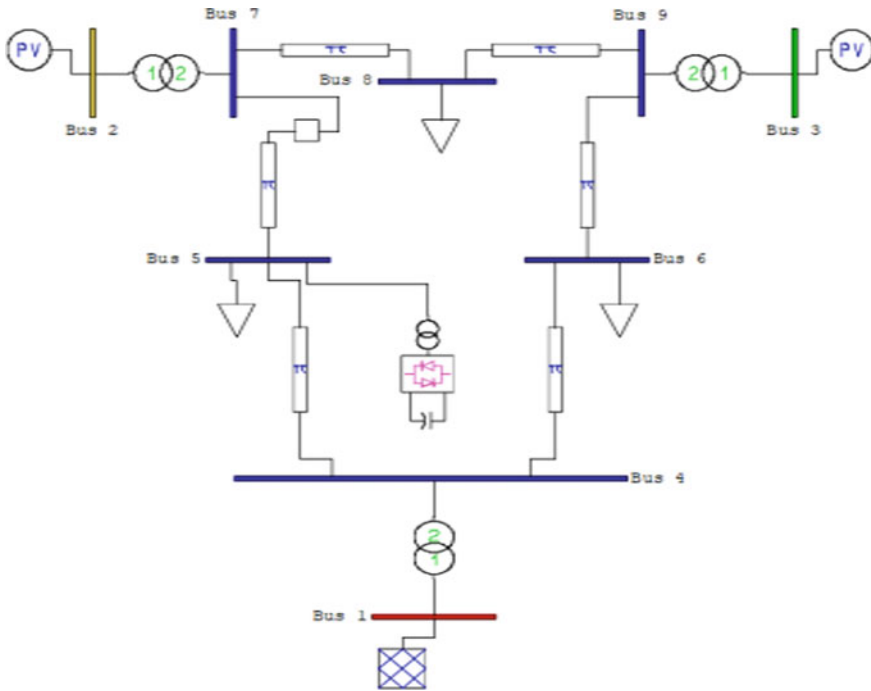


Fig. 4 9 bus IEEE network with STATCOM

from which maximum loading parameter is obtained. Detection of severe line or bus is carried out which results in unstable condition of the network.

6.1 Analysis Without Shunt FACTS Device

Voltages in per unit after line outage are shown in Table 3. Result shows that when the line between 7 and 8 gets out, the voltage at Bus 5 is very less as compared to other line outage. Voltages in per unit and voltage phase in radians, at various buses with line outage without using SVC and STATCOM, are shown in Fig. 6.

6.2 Analysis with STATCOM

Voltages in per unit after line outage with STATCOM are shown in Table 4. Result shows that the voltage at Bus 5 is increased as compared to other line outage without STATCOM. Bus voltages in per unit and voltage phase in radians, at various buses with line outage using STATCOM, are shown in Fig. 7.

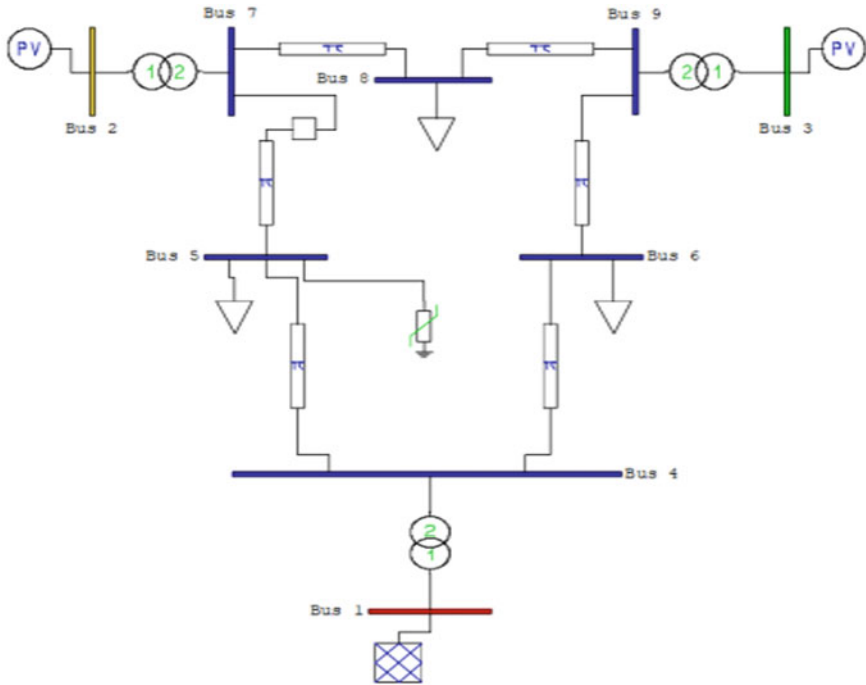


Fig. 5 9 bus IEEE network with SVC

Table 1 Generator, transformer and load data

S. No.	Machine	Rating
1	Generator 1	247.5 MW, 16.5 kV
2	Generator 2	192 MW, 18 kV
3	Generator 3	128 MW, 13.8 kV
4	Transformer 1	77 MW, 16.5/230 kV
5	Transformer 2	163 MW, 18/230 kV
6	Transformer 3	86 MW, 13.8/230 kV
7	Three phase series RLC load at bus 5, 6 and 8	134.62 MW, 94.86 MW, 105.94 MW

Table 2 Power line data

Bus	R (pu)	X (pu)	B (pu)
4 to 5	0.017	0.092	0.079
6 to 9	0.039	0.17	0.179
5 to 7	0.032	0.161	0.153
7 to 8	0.0085	0.072	0.0745
8 to 9	0.0119	0.1008	0.1045
4 to 6	0.017	0.092	0.079

Table 3 Outage bus voltages without shunt FACTS device

Bus (pu)	Voltages during line outage without FACTS device					
	7 to 8	6 to 9	8 to 9	4 to 5	5 to 7	4 to 6
Bus1	1.04	1.04	1.04	1.04	1.04	1.04
Bus2	1.025	1.025	1.025	1.025	1.025	1.025
Bus3	1.025	1.025	1.025	1.025	1.025	1.025
Bus4	0.820	0.782	0.868	1.030	0.886	0.9985
Bus5	0.536	0.551	0.674	0.584	0.735	0.9272
Bus6	0.773	0.614	0.800	0.998	0.814	0.692
Bus7	0.827	0.852	0.779	0.915	0.9718	0.975
Bus8	0.779	0.854	0.596	0.929	0.920	0.927
Bus9	0.907	0.965	0.973	1.001	0.955	0.944

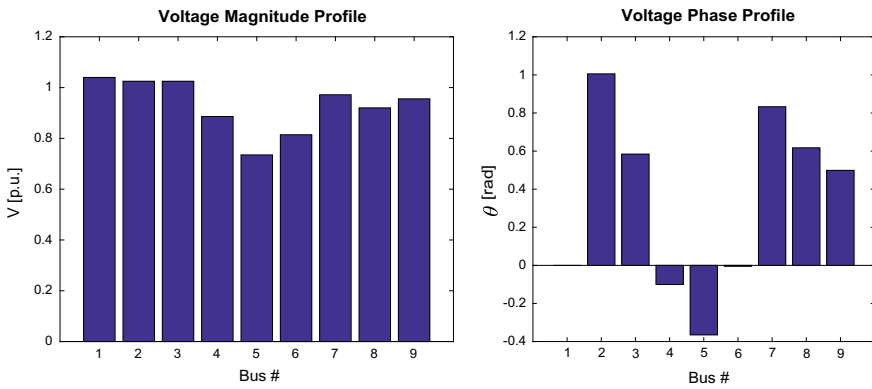


Fig. 6 Voltage magnitude profile and phase profile without shunt FACTS device

Table 4 Bus voltages after line outage with STATCOM

Bus (p.u)	Voltages during line outage with STATCOM					
	7 to 8	6 to 9	8 to 9	4 to 5	5 to 7	4 to 6
Bus1	1.04	1.04	1.04	1.04	1.04	1.04
Bus2	1.025	1.025	1.025	1.025	1.025	1.025
Bus3	1.025	1.025	1.025	1.025	1.025	1.025
Bus4	0.982	0.967	0.988	1.033	0.963	1.022
Bus5	0.944	0.937	0.960	0.831	0.921	0.986
Bus6	0.913	0.862	0.916	1.006	0.886	0.705
Bus7	0.986	0.972	0.939	0.979	0.980	0.990
Bus8	0.821	0.944	0.837	0.975	0.936	0.939
Bus9	0.946	0.998	1.006	1.017	0.975	0.951

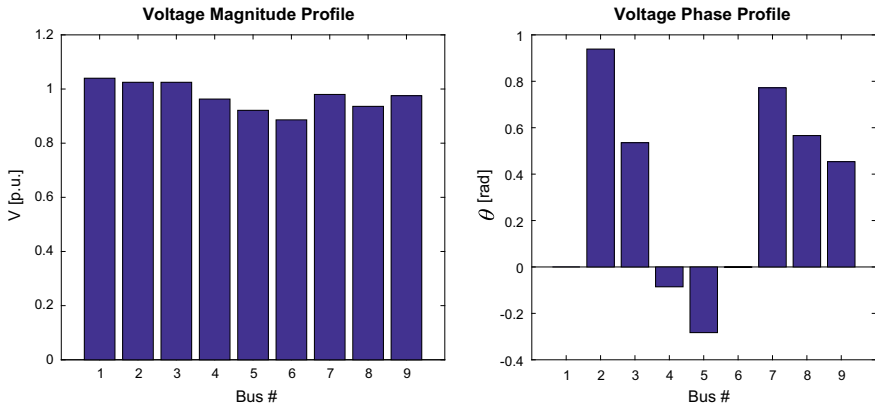


Fig. 7 Voltage magnitude profile and phase profile with STATCOM

Table 5 Bus voltages after line outage with SVC

Bus (p.u)	Voltages during line outage with SVC					
	7 to 8	6 to 9	8 to 9	4 to 5	5 to 7	4 to 6
Bus1	1.04	1.04	1.04	1.04	1.04	1.04
Bus2	1.025	1.025	1.025	1.025	1.025	1.025
Bus3	1.025	1.025	1.025	1.025	1.025	1.025
Bus4	0.833	0.972	0.965	1.030	0.964	1.025
Bus5	0.588	0.950	0.957	0.626	0.928	0.992
Bus6	0.769	0.864	0.8465	0.996	0.885	0.724
Bus7	0.838	0.974	0.8201	0.925	0.978	0.9940
Bus8	0.746	0.9442	0.597	0.934	0.933	0.9449
Bus9	0.895	0.998	0.980	1.002	0.974	0.9566

6.3 Analysis with SVC

Voltages in per unit after line outage with SVC are shown in Table 5. Result shows that the voltage at Bus 5 is increased as compared to other line outage without SVC. Bus voltages in per unit, voltage phase in radians, at various buses with line outage using SVC shown in Fig. 8.

7 Conclusions

In this paper, contingency analysis is done for grid interconnection network. STATCOM and SVC a shunt power device is used for improving performance of

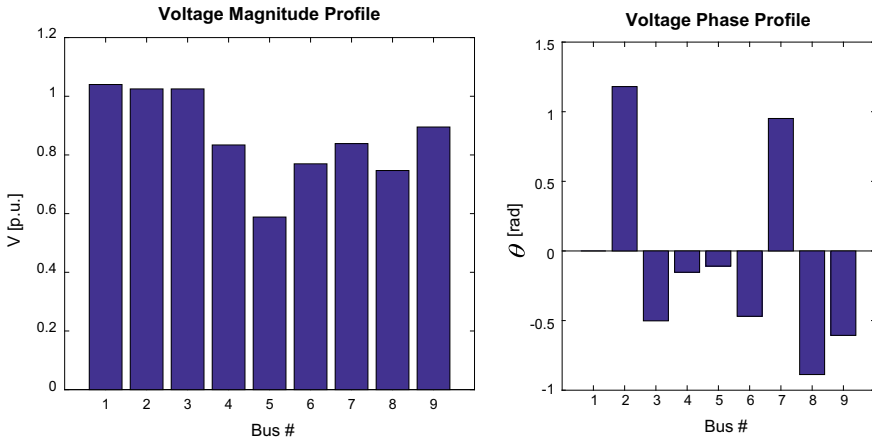


Fig. 8 Voltage magnitude profile and phase profile with SVC

power network during line outages. PSAT software is used for testing purpose. The contingency analysis using N-R method and continuation power flow is carried out for line outage. From which Bus 5 obtained as weakest bus as compared to other bus of the network. So SVC and STATCOM are connected at Bus 5 for performance enhancement of power network at contingency condition. Rating of power device is decided by considering loading parameter limit of original network. The result of 9 bus IEEE test network shows that by using STATCOM and SVC shunt power device increases the voltage level of bus and improves performance of the network at contingency condition.

References

1. P. Kundur, *Power Systems Stability and Control* (McGraw Hill Inc., New York, 1994)
2. B. Nitve, R. Naik, Steady state analysis of IEEE-6 Bus System Using PSAT power toolbox. *IJESIT* **3**, 197–203 (2014)
3. D. Asija, P. Choudekar, K.M. Soni, S.K. Sinha, Power flow study and contingency status of WSCC 9 bus test system using MATLAB, in *2015 International Conference on Recent Developments in Control, Automation and Power Engineering (RDCAPE)*, pp. 338–342
4. V.K. Chandrakar, A.G. Kothari, Fuzzy-based Static Synchronous Compensator (STATCOM) for improving transient stability performance. *Int. J. Energy Technol. Policy* **5**(6), 692–707 (2007)
5. M. Darabian, A. Jalilvand, Improving power system stability in the presence of wind farms using STATCOM and predictive control strategy. *IET Renew. Power Gener.* **12**(1), 98–111 (2018)
6. M.V. Wankhede, V.K. Chandrakar, Static synchronous compensator (STATCOM) for the improvement of the electrical system performance with Non Linear load. *Int. J. Power Syst.* **1**, 27–32 (2016)

7. V.K. Chandrakar, A.G. Kothari, Comparison of RBFN and Fuzzy based STATCOM controllers for transient stability improvement, in *International Conference on Electrical Machines and Power Electronics*, pp. 520–525 (2007)
8. V.K. Chandrakar, A.G. Kothari, Optimal location for line compensation by shunt connected FACTS controller, in *Fifth International Conference on PEDS*, vol. 1, pp. 151–156 (2003)
9. E. Barrios-Martinez, C. Ángeles-Camacho, Technical comparison of FACTS controllers in parallel connection. *J. Appl. Res. Technol.* **15**, 36–44 (2017)
10. S.K. Umathe, STATCOM: theoretical aspects & experimental study. *Int. J. Electr. Eng.* **4**(4), 439–448 (2011)
11. M. Sahraei-Ardakani, X. Li, P. Balasubramanian, K.W. Hedman, M. Abdi-Khorsand, Real-time contingency analysis with transmission switching on real power system data. *IEEE TRANS. Power Syst.* (2015)
12. Z. Li, J. Wang, H. Sun, Q. Guo, Transmission contingency analysis based on integrated transmission and distribution power flow in smart grid. *IEEE Trans. Power Syst.* (2014)
13. P. Domkawale, V.K. Chandrakar, Voltage stability enhancement in large power system by using STATCOM. *Int. Res. J. Eng. Technol.* **4**, 1–7 (2017)
14. M. Majidi-Qadikolai, R. Baldick, Integration of N-1 contingency analysis with systematic transmission capacity expansion planning: ERCOT case study. *IEEE Trans. Power Syst.* (2015)
15. F. Albuyeh, A. Bose, B. Heath, Reactive power considerations in automatic contingency selection. *IEEE Trans. Power Apparatus Syst.* **PAS-101**(1), 107–112 (1982)
16. S. Zonouz, C.M. Davis, K.R. Davis, R. Berthier, R.B. Bobba, W.H. Sanders, SOCCA: a security-oriented cyber-physical contingency analysis in power infrastructures. *IEEE Trans. Smart Grid* **5**(1) (2014)
17. V.P. Rajderkar, V.K. Chandrakar, Comparison of series FACTS devices via optimal location in a power system for congestion management, in *Asia-Pacific Power and Energy Engineering Conference*, pp. 1–5 (2009)
18. N.I. Yusoff, A.A. Mohd Zin, A. Bin Khairuddin, Congestion management in power system: a review, in *IEEE 2017, International Conference on Power Generation systems and Renewable Energy Technologies, Recent Advances in Power Systems*, pp. 425–433
19. V. Ajjarapu, *Computational Techniques for Voltage Stability Assessment and Control. Power Electronics and Power Systems* (Springer, 2007), p. 49
20. V.J. Mishra, M.D. Khardennis, Contingency analysis of power system, in *2012 IEEE Students' Conference on Electrical, Electronics and Computer Science*
21. M.K. Jalboub, A.M. Ithal, H.S. Rajamani, R.A. Abd-Alhameed, A.M. Ithal, Determination of static voltage stability-margin of the power system prior to voltage collapse, in *2011 8th International Multi-Conference on Systems, Signals and Devices (SSD)* (2011)

Comparative Analysis of Modified Firefly Algorithm with Firefly Algorithm MPPT for PV Module Under Partial Shaded Condition



S. Sachin, Anusha Vadde, and Sowmyashree

Abstract Photovoltaic (PV) modules will decrease their output power when they are subjected to partial shading conduction (PSC). Hence, there are different maximum power point tracking (MPPT) controlled algorithms designed to decrease the shaded effects on PV module. Firefly algorithm (FA) which is known as meta-heuristic algorithm is introduced, to get global maximum power (GMP) in PSC conduction. Even though FA can successfully trap GMP, there are some problems which affected the operation of MPPT. This paper helps to overcome the operational problem by comparing modified FA (MFA) with the existing FA. This MFA can reduce the number of total computation operation, and GMP time is converged for the present FA requirement. Simulation results exhibit that the proposed technique can track the global power (GP) under different PSC, the converging time is quick, and the voltage and power variations are reduced efficiently with MFA.

Keywords Maximum power point tracking (MPPT) · Maximum power point (MPP) · Partial shading (PS) · Partial shading conduction (PSC) · Photovoltaic module (PVM) · Firefly algorithm (FA) · Modified firefly algorithm (MFA) · Global maximum power (GMP) · Local maximum power (LMP) · Continuous conduction mode (CCM) · Global power (GP) · Single-ended primary-inductor converter (SEPIC)

1 Introduction

In PV module, the output power keeps changing with change in the environmental conditions such as irradiation of the sun and temperature of the cell due to which the efficiency of the PV module is reduced [1]. During non-sunny days, the effect of irradiation of sun will be lower which results in PSC effect. The shaded cells in a PVM work as a load; when it is reverse bias, it leads to hot spot in PV cells. To reduce this effect, the generated current by a non-shaded cell will be conducted through bypass

S. Sachin · A. Vadde (✉) · Sowmyashree
Ramaiah University of Applied Science, Bangalore, India

diodes in a PV module [2]. Overall, the bypass diode will alter the power, voltage and current characteristics of (P–V) and (I–V). To get effective operation of PVM, it is necessary to track GMP. A converter is required to enhance and controls the voltage and power from PV module.

A DC to DC converter is used to increase the voltage at the output. The switching operation of the converter is mainly dependent on values of inductance and capacitance. SEPIC converter has more effective switching operation when compared with Buck–Boost converter and Cuk converter [3].

To regulate voltage, power and current the perturbation method is used. The difficulty in this method is to differentiate between LMP and GMP. Due to LMP, power loss is high. The implementation of firefly for MPPT under PSC results in better performance than perturbation & observation method (P&O). FA leads to decrease in the efficiency of the PVM; for every iteration, the computation time is more, and if the movement of fly increases, then the brighter fly will increase. Because of this, the tracing time/oscillating time will be excessively high for GMP. Hence, this paper will improve the existing FA by reducing the tracing time/oscillating time and enhance the overall system efficiency.

2 Block Diagram

The proposed system which includes PV module, SEPIC converter, MPPT controller and a load is connected as shown in Fig. 1. The PVM is connected to the SEPIC converter, to meet the load requirement. The FA is modified to track maximum power from the PVM and also reduce the oscillation time by which the efficiency can be increased.

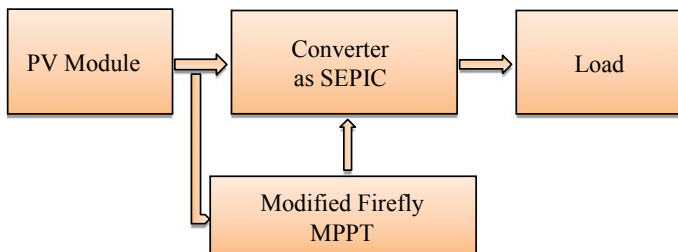


Fig. 1 Block diagram of the designed system

2.1 Design of Photovoltaic Cell

The solar panel is designed for 20 W. The modeling is based on the electrical characteristics of the panels and provides information about the operational state of the system. The voltage of 36 V and a current of 0.66 A are developed in a photovoltaic cell.

2.2 Design of SEPIC Converter

Peak-to-peak ripple current at minimum voltage of approximately 40% of the average input power is considered. Inductor input current is determined using Eq. (1).

$$\Delta I_l = I_{\text{out}} \times \frac{V_{\text{out}}}{V_{\text{in}(\text{min})}} \times 40\% \quad (1)$$

L_1 and L_2 are chosen as inductor ripple current, and the peak ripple currents of L_1 and L_2 are computed using Eqs. (2) and (3).

$$I_{L1(p)} = I_{\text{out}} \times \frac{V_{\text{out}} + V_D}{V_{\text{in}(\text{min})}} \times \left(1 + \frac{40\%}{2}\right) \quad (2)$$

$$I_{L2(\text{peak})} = I_{\text{out}} \times \left(1 + \frac{40\%}{2}\right) \quad (3)$$

The forward voltage drop of the diode is considered as V_D , and SEPIC converter is operating in CCM. The maximum and minimum duty cycles are stated in Eqs. (4) and (5).

$$D_{\text{min}} = \frac{V_{\text{out}} + V_D}{V_{\text{in}(\text{max})} + V_{\text{out}} + V_D} \quad (4)$$

$$D_{\text{max}} = \frac{V_{\text{out}} + V_D}{V_{\text{in}(\text{min})} + V_{\text{out}} + V_D} \quad (5)$$

Power metal-oxide-silicon-field effect transistor (MOSFET) selection is based on MOSFET peak current which is in Eq. (6) and root mean square (RMS) current through switch is in Eq. (7).

$$I_{Q1(p)} = I_{L1(p)} + I_{L2(p)} \quad (6)$$

$$I_{Q1(p)} = I_{\text{out}} \sqrt{\frac{(V_{\text{out}} + V_{\text{in}(\text{min})} + V_D) \times (V_{\text{out}} + V_D)}{V_{\text{in}(\text{min})}^2}} \quad (7)$$

Table 1 SEPIC converter specification

Parameter	Characteristics	Specification
V_{in}	Input voltage	36–110 V
V_o	Output voltage	110 V
L_1	Inductor	12.3 mH
C_1	Capacitor	0.01 mF
R_L	Load resistor	605 Ω
F	Frequency	10 kHz
P	Power	20 W

Selection of output capacitor is based on the capacitor output RMS current which is,

$$I_{\text{cout(rms)}} = I_{\text{cs(rms)}}$$

Peak-to-peak ripple is assumed as 2% of equivalent series resistance (ESR). The capacitor at the output end is in Eq. (9).

$$\text{ESR}_L \leq \frac{V_{\text{ripple}} \times 0.5}{I_{L1(p)} + I_{L2(p)}} \quad (8)$$

Capacitor is

$$C_{\text{out}} \geq \frac{I_{\text{out}} \times D_{\text{max}}}{V_{\text{ripple}} \times 0.5 \times f_{\text{sw}}} \quad (9)$$

Input capacitor selection:

$$I_{C_{\text{in}}(\text{rms})} = \frac{\Delta I_L}{\sqrt{12}} \quad (10)$$

Output diode selection:

The measured reverse voltage should be greater than sum of input and output voltage ($V_{\text{in}} + V_{\text{out}}$), and diode average current at maximum charge is always equal to the output current (Table 1).

2.3 Proposed Algorithms

The process of firefly and modified firefly algorithms are explained in flowcharts. Firefly, MPPT algorithm has been used for simulation. The effectiveness can be figured by tracking the proportion between yield power and most extreme intensity of the PV array under steady-state conditions [3].

There are two sensor hubs, one is yield voltage sensor to detect the yield voltage and the other is yield current sensor to detect the yield current of PV array. Detection of current and voltage happens simultaneously based on listed conditions:

- $\frac{\Delta I}{\Delta V} > 0$ The voltage is less than MPPT (left of MPPT)
- $\frac{\Delta I}{\Delta V} = 0$ PV voltage = MPPT Voltage
- $\frac{\Delta I}{\Delta V} < 0$ The PV voltage is greater than MPPT voltage (right of MPPT).

Here, I and V refer to PV current and voltage, respectively. There are three factors that predominantly influence the execution of any PV system which are sun powered light, surrounding temperature and load conditions [4]. The drawback of this method is the MPP is lost, due to the change in isolation level based on the solar irradiation.

The progression with that PVM yield voltage V_{PV} and yield current I_{PV} is recognized. At that point, power is determined as $P(n)$ and the estimated value of the past example is $P(n - 1)$ yielding ΔP . If the MPP is $\frac{dP}{dV} > 0$ or $\frac{dP}{dV} < 0$, working point is moving ceaselessly from MPP as shown in Fig. 2. The drawback of P&O algorithmic program the voltage of MPP is half tracked to fulfill $(\frac{dP}{dV})_{PV} = 0$ [5].

Firefly Algorithm

The PV voltage is the managed variable, which is simple to actualize as the MPPT controller can rapidly choose the underlying focuses [6], according to the level of open-circuit voltage (V_{oc}). Obligation cycle is likewise a controlled variable, so it should likewise be possible along these lines.

The FA strategy has three major assumptions. Initially, firefly will always move toward brighter side, and they are genderless also progressively appealing in single until that every one of them had been thought about. Also, the engaging quality of a firefly relies upon their brilliance, which is identified with the separation among itself and different flies [7–11]. Be that as it may, the separation between them

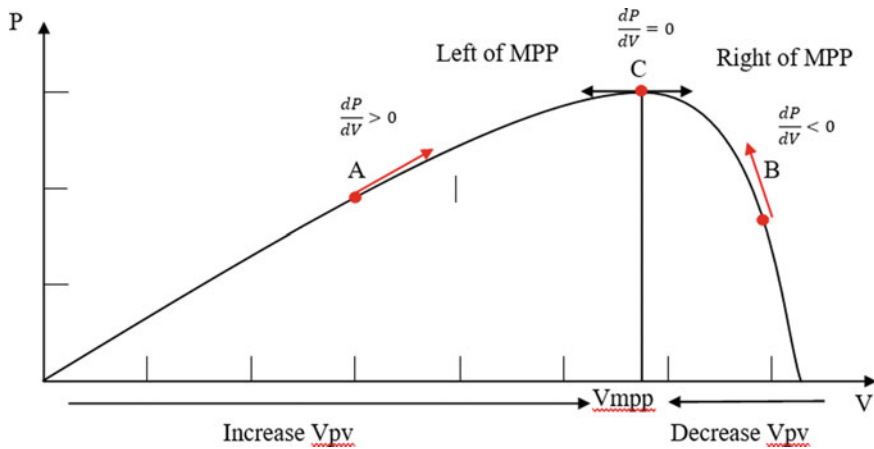


Fig. 2 Power–voltage (P–V) graph of a MPPT

will increment because of which flies move away the air. $P_{G_{best}}$ is refreshed by positioning the fireflies. From the target of the given capacity beneath the light force of firefly is resolved. Scientifically, the FA calculation can be communicated by three conductions. ' β ' is given by the expressed condition.

$$\beta(r) = \beta_0 \exp(-\Upsilon r^m), m \geq 1 \quad (11)$$

where β_0 is the underlying allure when r is equal to zero, where r is the separation of two fireflies, Gamma (γ) is an ingestion for the coefficient to control the decrease of light power, and whole number is given by ' m '. The assessment of separation of two flies i and j , which are positioned at x_i and x_j , is characterized by Euclidean separation.

$$r_{ij} = \|x_i - x_j\| = \sqrt{\sum_k^d (x_{i,k} - x_{j,k})^2} \quad (12)$$

where x_i, k and x_j, k are the k th part of the spatial directions of the i th and j th firefly, respectively. The quantity of measurements is d . The resolved of a development of firefly is in Eq. (13).

$$x_i = x_i + \beta_0 e^{-\Upsilon r_{ij}^2} (x_j - x_i) + \alpha \cdot \left(\text{randn} \frac{1}{2} \right) \quad (13)$$

where $\alpha \in [0, 1]$ and randn is an irregular bother esteem. Equation (13) explains that the inadvertent of more brilliant firefly will be influenced by the development of fly. The inadvertent gives decent method for the fly to go far from neighborhood for searching the pursuing for worldwide calculation. By using MPPT, the value of V_{ref} can be found. In this manner, the status of (x_i) in Eq. (13), is determined. The V_{ref} is dependent on the 2nd and 3rd term from Eq. (13).

The fireflies are instated and assessed in a progressive way on the grounds that the converter can react to just each order in turn. When V_{ref} is applied, and N is initialized to the quantity of firefly, the positive number is N . In additional, to the worldwide most extreme force, $P_{G_{best}}$ is instated by the force level and the positioning will be given by it. Note that the force is a result of estimated voltage (V_{pv}) and current (I_{pv}). At that point, the code will continue to the settled form of circles. The internal circle has a power of the fly I is a lot of irregular worth. This settles circle will be continued till all flies are looked with each other. From the instance MPPT, ' d ' will be equal to 1

$$r_{ij} = \sqrt{(x_i - x_j)^2} \quad (14)$$

The $P_{G_{best}}$ is refreshed by positioning the fireflies. The terminology used in flowchart is shown in Fig. 3 according to Table 2.

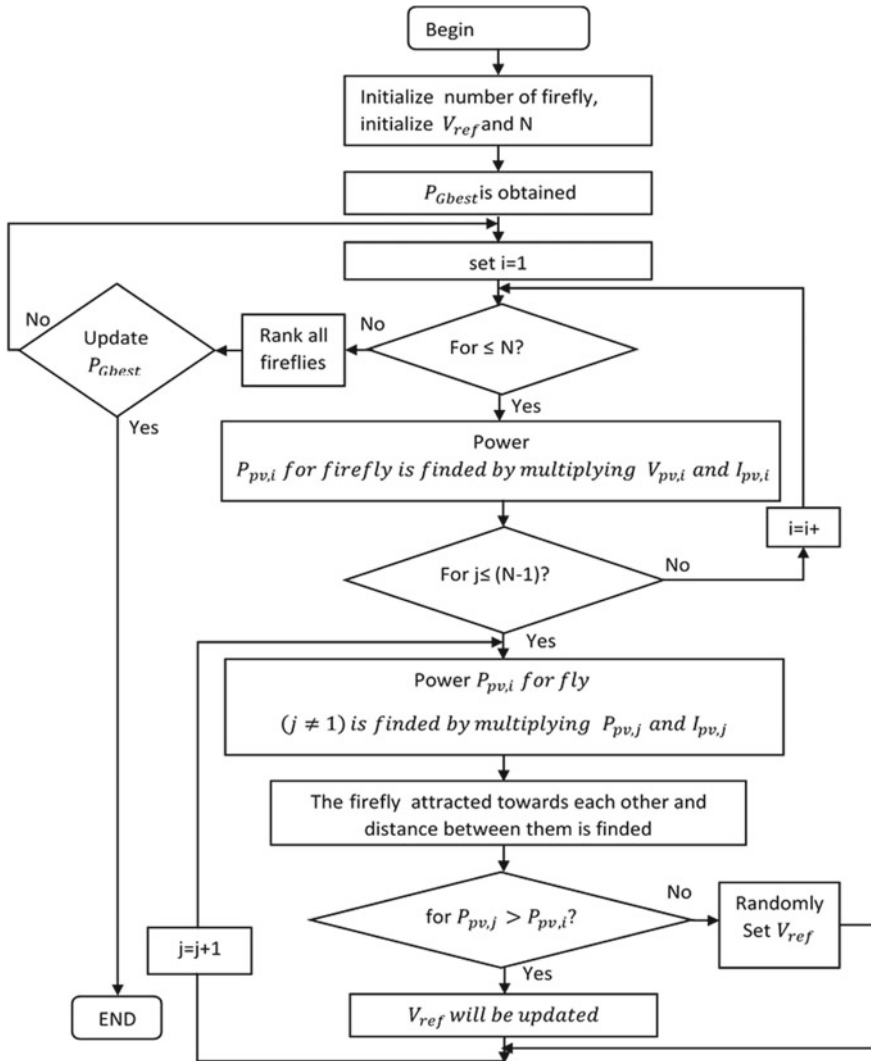


Fig. 3 Flowchart of firefly algorithm

Table 2 Terms used in PV module and FA algorithm

PV module	FA algorithm
Reference voltage (V_{ref})	Position of firefly
Difference in voltage (ΔV_{ref})	Distance between fly
Exponential function of ΔV_{ref}	Attractiveness of fly
Power of module (P_{pv})	Brightness of firefly
Global maximum power (P_{Gbest})	Brightest firefly

Modified Firefly Algorithm

Toward the brightness of firefly, the position of each firefly will change. This is due to comparison of firefly with each other and has a movement; this is shown by Eq. (13), and this is the biggest drawback of FA.

The MFA flowchart is similar to the present FA flowchart. The program is initialized to form a loop. The fly_{*i*} ($P_{pv,i}$) and fly_{*j*} ($P_{pv,j}$) are the intensities of flies which are calculated. In the first loop, $P_{pv,j}$ and $P_{pv,i}$ are compared with each other for finding which fly is brighter and L is the number of flies. The updated x_i is the brighter fly than other fly. With respect to average value, the x_i will be updated. According to the modified firefly flowchart shown in Fig. 4, it finds the average of all firefly, and then, each fly will be compared with each other. Hence, the number of computations are decreased compared to FA.

3 Characteristics of the Proposed System

Switching characteristics of SEPIC converter, oscillating time and efficiency plots for the system are detailed in Sects. 3.1, 3.2 and 3.3.

3.1 Switching Characteristic of SEPIC Converter

A DC-DC converter has their own switching characteristics. The factors such as input pulse, inductor current, output current and voltage across capacitor affect the switching operation of the converter.

A SEPIC converter is built with an input and output inductance of 12.3 mH and capacitance of 0.01 mF for a load of 605 Ω . The input pulse for the SEPIC is given for 0.0001 s. The first and second inductor currents are 1.1 A and -0.1 A, respectively, and capacitance switching voltage is 30 V for the input voltage of 36 V. The switching characteristics are as shown in Fig. 5.

3.2 Characteristics of Firefly Algorithm

Firefly algorithm is designed for increasing the efficiency of the PV module which is designed for the input voltage of 36 V as shown in Fig. 6. The output voltage is reduced at 0.25 s and maintains 90 V further simulation time as shown in Fig. 7 and the output current of 0.17 A as shown in Fig. 8. The system will be under damp condition; hence, the time taken for oscillation will be in the range of 0.25–0.33 s. Therefore, settling time is more; hence, the overall efficiency of the system is around 89% as shown in Fig. 9.

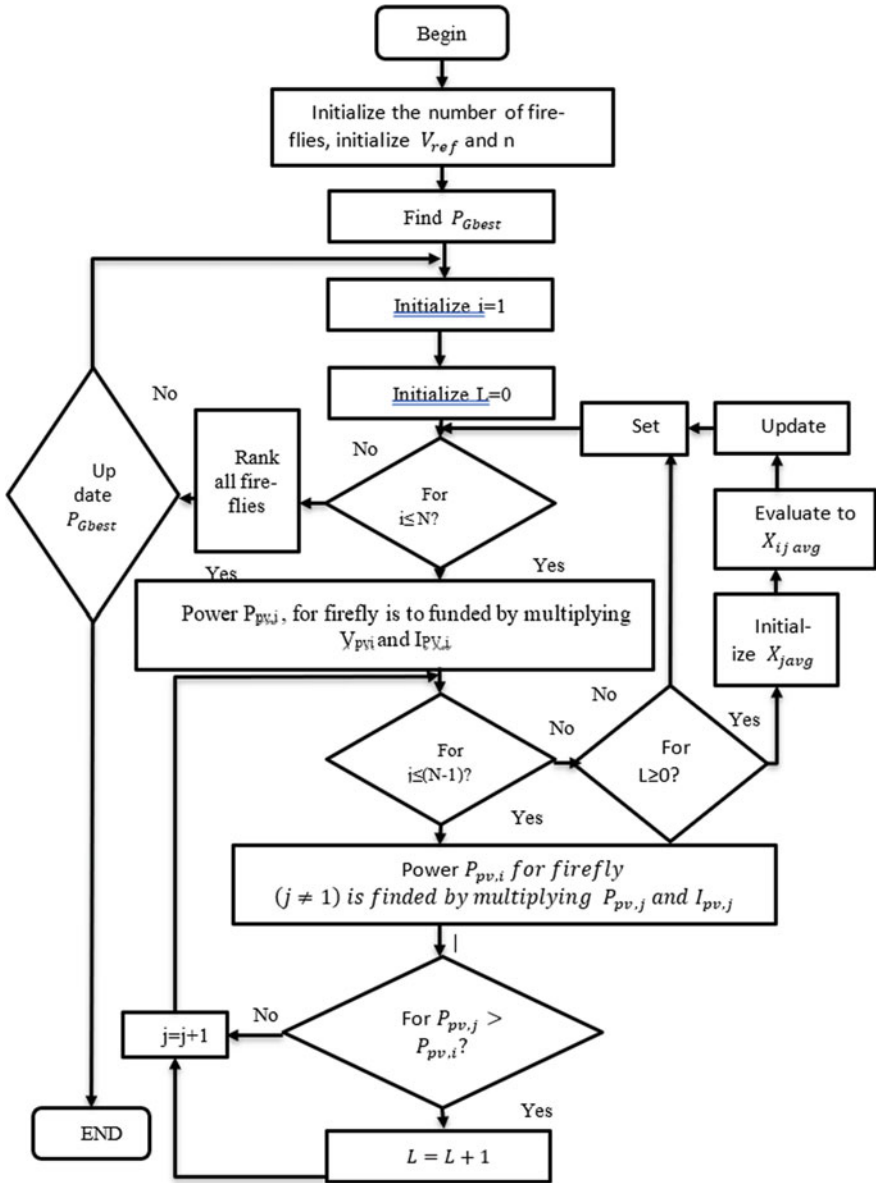


Fig. 4 Flowchart of modified firefly algorithm

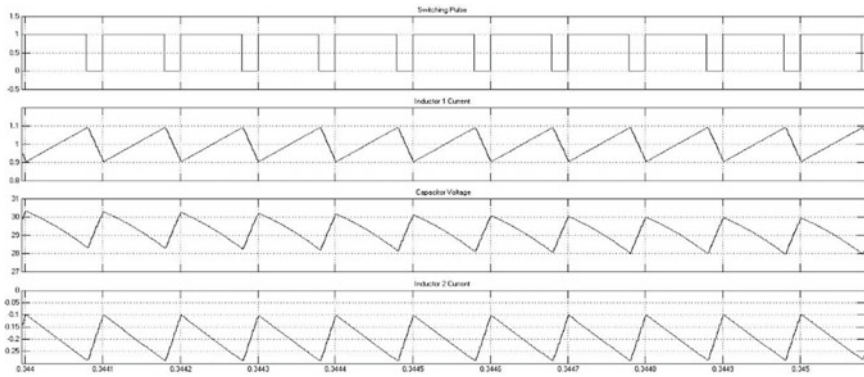


Fig. 5 Switching characteristics of SEPIC

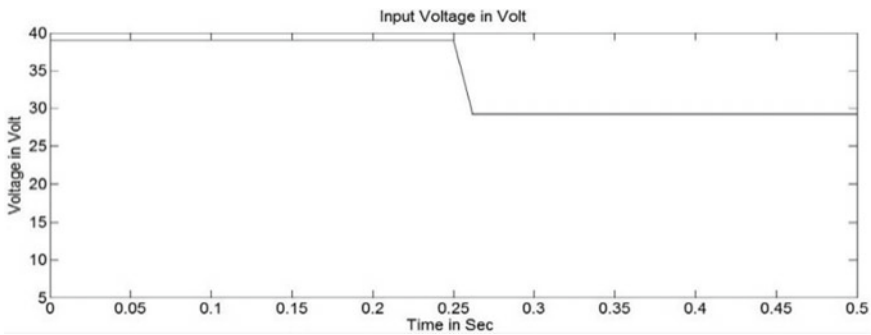


Fig. 6 Input voltage characteristics for FA

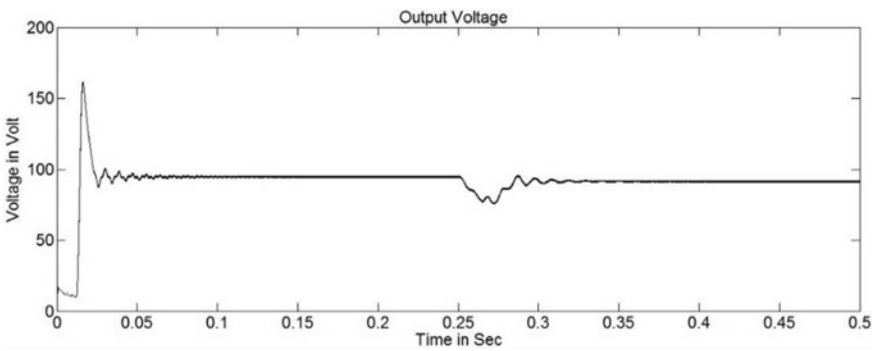


Fig. 7 Output voltage characteristics for FA

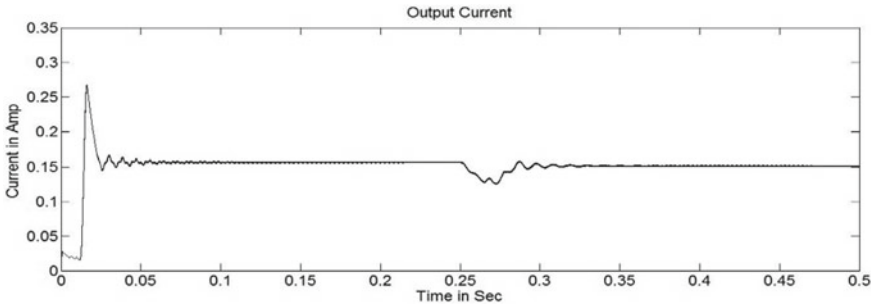


Fig. 8 Output current characteristics for FA

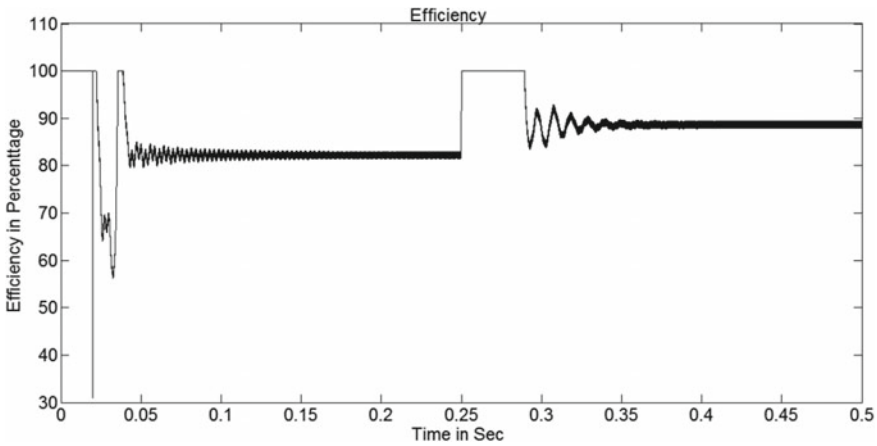


Fig. 9 System efficiency of FA

3.3 Characteristics of Modified Firefly Algorithm

Modified firefly algorithm is designed such that the number of iterations is reduced for the same input value. Hence, by which, the oscillation time is reduced leading to increase in overall efficiency of the system as shown in Fig. 13. To achieve the maximum efficiency from MFA, the average of all firefly at a time will lead to reduction in iteration. The oscillation time taken for settling is around 0.25–0.29 s. Hence, the overall system efficiency is increased. For the given input voltage of 36 V, the output will be 110 V and the output current is 0.18 A as shown in Figs. 10 and 11, respectively. Figure 12 represents the input and output characteristics of modified firefly algorithm.

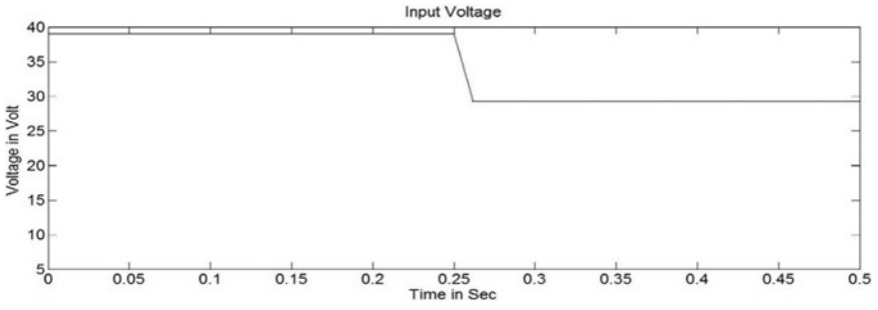


Fig. 10 Input voltage characteristics for MFA

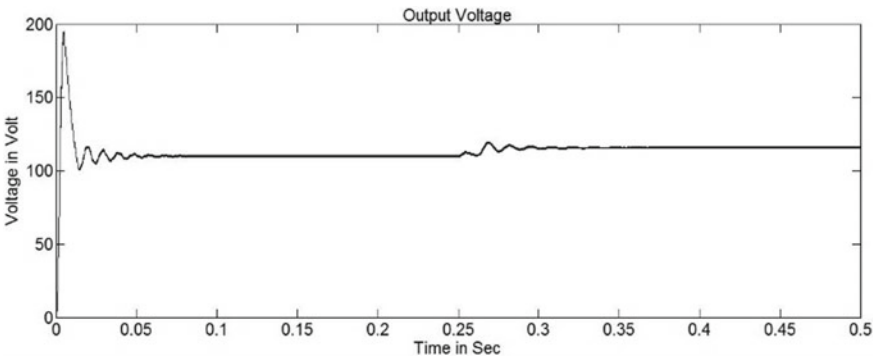


Fig. 11 Output voltage characteristics for MFA

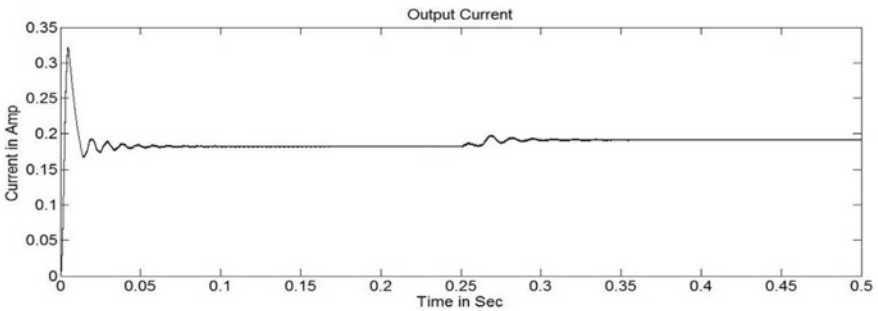


Fig. 12 Output current characteristics for MFA

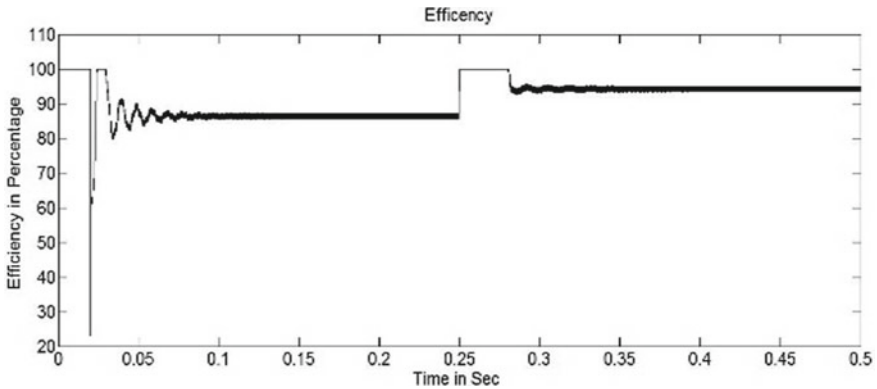


Fig. 13 System efficiency of MFA

Table 3 Comparison of oscillation time for FA and MFA algorithms of SEPIC converter

Types of algorithm applied for SEPIC converter	Oscillation time in s	Overall system efficiency (%)
Firefly	0.25–0.29	89
Modified firefly	0.25–0.27	96

4 Conclusions

This paper analyzes and compares MPPT system to follow the GMP for FA and MFA. The methodology presents calculation and utilization and normalizes the fireflies toward the point while not meandering toward all the more brilliant flies. The consequences of all the calculation procedures are given in Table 3.

The results are shows that the proposed firefly algorithm is better than firefly calculation in oscillation time and furthermore decline the amount of tasks per cycle with referred to Table 3. Modified firefly calculation has more proficiency than firefly calculation where occurrence of quick light changes.

References

1. W.B.F. Corporation, *Utility Scale Solar Power Plants: A Guide for Developers and Investors* (World Bank, India, 2012).
2. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
3. R.C.N. Pilawa-Podgurski, W. Li, I. Celanovic, D.J. Perreault, Integrated CMOS energy harvesting converter with digital maximum power point tracking for a portable thermophotovoltaic power generator. *IEEE J. Emerg. Selected Topics Power Electron.* **3**(4), 1021–1035 (2015)

4. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol 661, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
5. H. Patel, V. Agarwal, Maximum power point tracking scheme for PV systems operating under partially shaded conditions. *IEEE Trans. Ind. Electron.* **55**(4), 1689–1698 (2008)
6. I. Abdalla, J. Corda, L. Zhang, Multilevel dc-link inverter and control algorithm to overcome the PV partial shading. *IEEE Trans. Power Electron.* **28**(1), 14–18 (2013)
7. D. Nguyen, B. Lehman, An adaptive solar photovoltaic array using model-based reconfiguration algorithm. *IEEE Trans. Industr. Electron.* **55**(7), 2644–2654 (2008)
8. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar pv arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON)* held at Deenbandhu Chhotu Ram University of Science and Technology, SONEPAT, India on Feb 28–March 1, 2020
9. Syafaruddin, E. Karatepe, T. Hiyama, Artificial neural network-polar coordinated fuzzy controller based maximum power point tracking control under partially shaded conditions. *IET Renew. Power Gener.* **3**(2), 239–253 (2009)
10. Y.-H. Liu, S.-C. Huang, J.-W. Huang, W.-C. Liang, A particle swarm optimization-based maximum power point tracking algorithm for PV systems operating under partially shaded conditions. *IEEE Trans. Energy Convers.* **27**(4), 1027–1035 (2012)
11. L.L. Jiang, D.L. Maskell, J.C. Patra, A novel ant colony optimization-based maximum power point tracking for photovoltaic systems under partial shaded conditions. *Energy Buildings* **58**(1), 227–236 (2013)

Active Power Control for Single-Phase Grid Connected Transformerless Inverter Photovoltaic System



Janardhan Gurram , G. N. Srinivas, and N. N. V. Surendra Babu

Abstract The level of increase in the trend of power conversion technology has made unbelievable dependency on renewable energy sources like solar, Wind etc. Besides the availability of abundant resources, increased efforts to harness the power out of wide variety inputs is phenomenal. Solar PV stands in first row now and for years to come, due to its availability at substantial cheaper cost, less maintenance, and a competitive efficiency over past few decades. Transformerless inverters for grid connected system are found to be beneficial in terms of space, cost, and efficiency constraints of the inverter. Though an inverter with transformer will provides galvanic isolation which can be made possible without transformers by decoupling ac and dc side of the inverter at regular intervals by incorporating extra switches on either side of the inverter which will be considered as different topology of inverter, subsequently by applying different switching modulation techniques. The paper considers the task of active power control in grid connected transformerless inverters using Highly Efficient Reliable Inverter Concept (HERIC) inverter to transfer the power from DC side of the inverter to AC side i.e. grid using MATLAB/SIMULINK. It also gives an impact in realizing the control strategy of injecting active power to grid. It emphasizes on filter requirements for grid connected PV system to feed sinusoidal current to grid by reducing unwanted harmonics into grid.

Keywords Transformerless inverter · HERIC · Grid connected inverter

J. Gurram (✉)

Department of EEE, CVR College of Engineering, Hyderabad, India

G. N. Srinivas

Department of EEE, JNTUH, Hyderabad, India

N. N. V. Surendra Babu

KIOT, Wollo University, Dessie, Ethiopia

e-mail: surendrababu@ieee.org

1 Introduction

The ever-increasing demand for power over past few decades, has made researchers and industry to concentrate on alternatives of energy resources. Continuous evolution in technology took the research to a new pace [1]. Adding to the abundant natural resources like solar, wind, biomass, etc. Efficient power conversion techniques, raise the curtains to harness the power from renewable energy sources. The dominance of Solar Photo voltaic (PV) has shown the path to renewable energy market [2]. This is because, solar PV provides clean energy at increasingly lower prices, less maintenance and high efficiency when compared with its counter parts in the renewable energy. The cost of solar PV is decreased due to its production in large scale [2]. The PV produces direct current i.e. DC, which is to be converted into alternate current i.e. AC to feed it to the utility grid. General grid connected inverter include either a low frequency transformer on AC side of the inverter or a high frequency transformer on DC side of the inverter which facilitates galvanic isolation between DC side and AC side of the inverter [3]. However, inclusion of transformer makes the system expensive, occupies larger space, causes increase in losses there by reduction in efficiency. Transformer can be avoided to overcome all the disadvantages that listed in inverters with transformer, but at the cost of galvanic isolation [4–6]. Due to absence of transformer, a leakage current is made to flow from PV module to the system, which is caused by the existence parasitic capacitance between PV module and ground [4]. The harmonics in grid current will be increased by the rise in leakage current, which will further create electromagnetic interference.

Standards are framed for maximum permissible leakage currents that can be allowed is 10 mA beyond which grid must be disintegrated from PV with in 0.3 s. as shown in Table 1 [7].

The bridge inverter with half wave output can remove leakage currents, but it requires double the input voltage than that is required in case of a full bridge inverter for the same output voltage of 230 V AC [8].

Table 1 Leakage current-grid disintegration

Leakage current in A	Grid disintegration time in sec
0.03	0.3
0.06	0.06
0.1	0.14

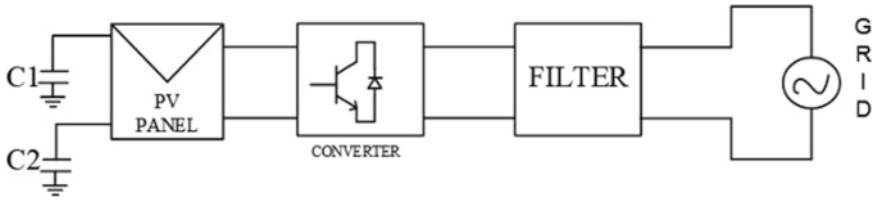


Fig. 1 Transformerless inverter

2 Grid Connected Inverters

2.1 Transformerless Inverters

With the advent of transformerless inverters, there has been a remarkable progress in research. A schematic of transformerless inverter is shown in Fig. 1, which indicates PV is connected to grid without a transformer [9]. A filter is connected between inverter and grid to obstruct the noise that may enter the grid. C_1 , C_2 are parasitic capacitance formed between PV module and ground [10]. The leakage current in grid connected transformerless inverters can be eliminated by employing various inverter topologies like H5, H6, HERIC improved H6 etc. [11]. The problem of leakage currents can also be addressed by applying variety of pulse width modulation (PWM) schemes in addition to the use of different inverter topologies. Bipolar sinusoidal pulse width modulation can reduce leakage current by maintaining constant common mode voltage. However, due to high switching losses, efficiency of such inverter is low. To counter this problem of leakage currents, many inverter topologies are presented in literature. Though literature includes the classification based on multiple factors, H-bridge transformerless inverters broadly categorized based on decoupling strategy i.e. AC side or DC side of the inverter [12]. H5, H6 inverter topologies will use DC side decoupling, whereas HERIC falls under AC side decoupling [13, 14].

In this paper HERIC inverter is used to analyze active power control into grid and filtering requirements.

2.2 HERIC Inverter

This topology shown in Fig. 2, combines the advantage of three level output voltage in h bridge inverter with unipolar pulse width modulation with reduced variation in common mode voltage in the inverter that uses bipolar pulse width modulation [15, 16]. Thus, without much effect on common mode voltage, efficiency can be improved.

During positive half cycle of load or grid voltage T6 is turned on T1 and T4 switched on at same time with switching frequency giving rise to positive output

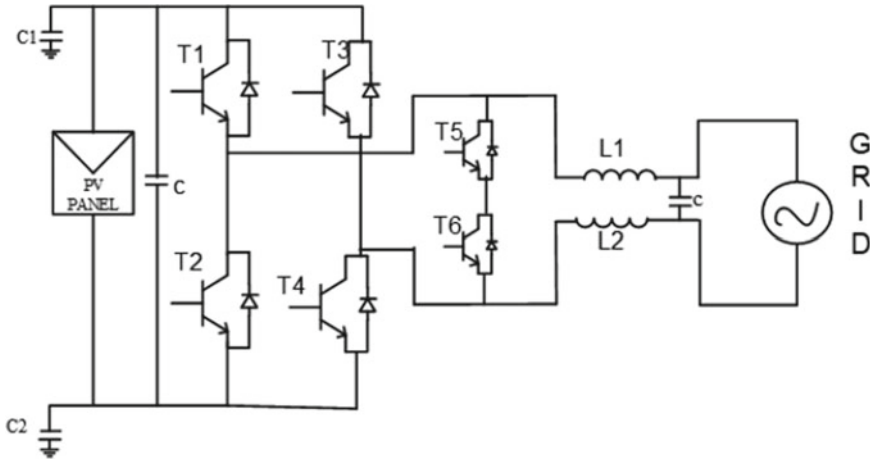


Fig. 2 Grid connected HERIC inverter

voltage while making T2, T3 are in off state. During active vector, PV supplies power to grid, whereas in zero state T1, T4 are in off state making the current freewheels through T6 and D5. During negative half cycle of grid voltage, T6 made off by turning on T5. In active vector state of negative part of the cycle T2, T3 are on to feed current from PV to grid. In freewheeling mode, i.e. zero state vector, current flows through D6, T5. Depending on positive or negative half cycle, by turning on either T5 or T6 and keeping T1 through T4 in off, thus making the output of the inverter shorted and isolated from PV. L_1 , L_2 and C in the circuit are filter inductance and capacitor, respectively [17, 18].

3 Principle of Power Control

The main function of grid connected PV system is to inject active power to grid. In addition to active power control, the control scheme gives the intense idea of reactive power control. In grid connected PV system power control is done by varying phase angle δ between inverter output voltage V_{inv} and grid voltage V_{grid} as shown in Fig. 3 [19, 20].

By varying phase angle ϕ between inverter output voltage and inverter output current I , variation over the power factor can be achieved, which subsequently attains the control over, active and reactive power as shown in Fig. 4 [19]. The active power injected into the grid by the inverter for a fundamental sinusoidal wave is given by the following equations [7].

$$P = \frac{V_{grid}}{\omega L} \cdot (V_m \cdot \sin \delta) \quad (1)$$

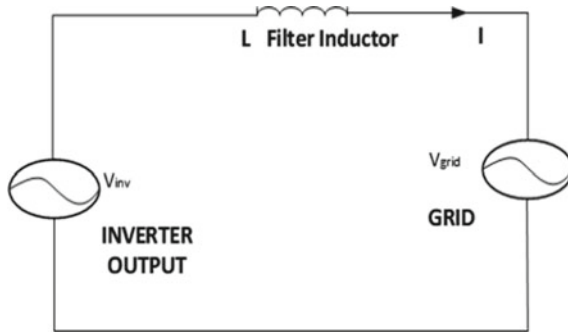


Fig. 3 Equivalent circuit of grid connected inverter

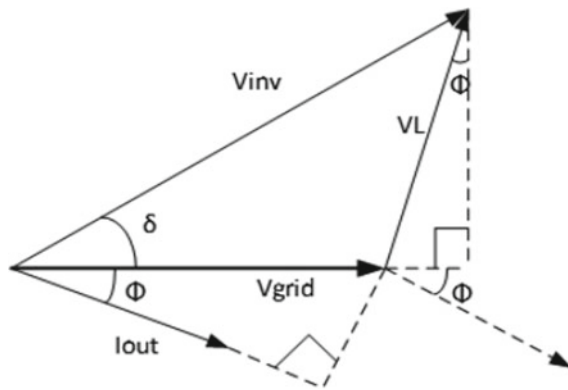


Fig. 4 Phasor representation of inverter connected to grid

$$Q = \frac{V_{grid}}{\omega L} \cdot (V_m \cdot \cos(\delta) - V_{grid}) \tag{2}$$

$$PF = \cos \phi = \frac{m \cdot V_{dc}}{\omega L} \sin(\delta) \tag{3}$$

Power factor is determined as shown in the Eq. (3) where V_{dc} is dc link voltage, m is the modulation index. Equation (2) indicates that inverter voltage magnitude influences the control of reactive power [4].

3.1 Control Implementation

To feed active current to grid from an external source like PV which is inherently a DC supply, it is required to analyze grid parameters and inverter output parameters

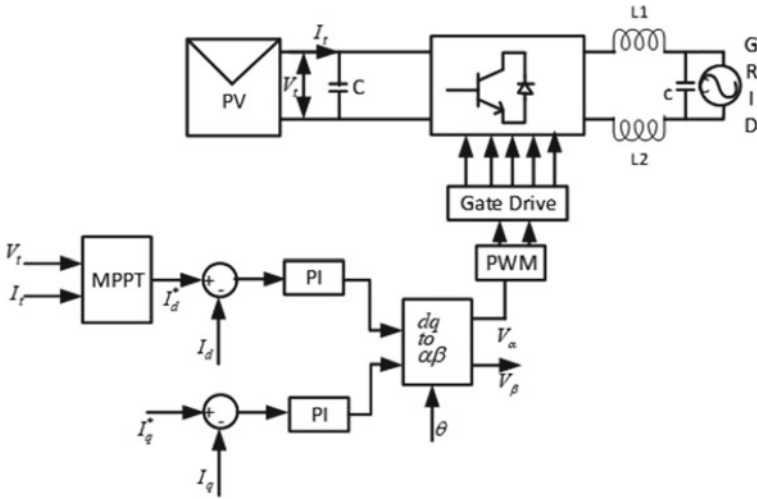


Fig. 5 Block diagram of $d-q$ axis control

like voltage, frequency, and phase angle [13]. It is required to control phase angle between output voltage of inverter and grid voltage to feed active power to grid. To supply reactive power to grid a distributed static compensator with a converter circuit is used whose voltage magnitude of the inverter is varied. In either of the cases, of controlling active or reactive power to grid, a decoupled control of d-axis and q-axis components of currents is employed and are controlled separately as shown in Fig. 5 [12].

The voltage and current from the PV are fed to the maximum power point tracking (MPPT) converter, whose output will generate current which is compared with a reference value and is passed through proportional integral (PI) controller that in turn generates voltage vector and is fed as one of the inputs to the block that converts rotatory state of reference frame to stationary frame of reference frame i.e. V_d-V_q to $\alpha\beta$ transformation [9]. Similarly, reactive component will generate a proportional voltage vector as the second input to the block.

Phase angle to be fed to the transformation block is derived from phase locked loop (PLL) as shown in Fig. 6a. Generation of reference current I_d, I_q derived from

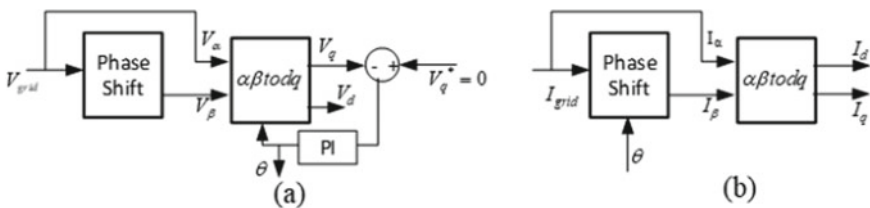


Fig. 6 a Determination of phase angle. b Reference currents generation

grid also shown in Fig. 6b. In Fig. 5, the inverter is of HERIC type whose pwm control signals are generated by comparing stationary reference frame sinusoidal voltage signal with that of high frequency carrier signal. To provide orthogonality in single phase currents a phase shift of $\pi/2$ radians is created with the help of a delay function [7, 9, 10].

$$V_{\text{grid}} = V \cdot \sin \omega t \quad (4)$$

$$I_{\text{grid}} = V \cdot \sin (\omega t + \theta) \quad (5)$$

Grid voltage or current will generate a orthogonal vector to represent $\alpha\beta$ frame of reference. Clarke's transformation matrix is used to convert stationary vectors $\alpha\beta$ to synchronously rotating reference vectors $d-q$. Clarke transformation matrix is given by Eq. (6).

$$\begin{bmatrix} d \\ q \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \cos(\theta) & -\sin(\theta) \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (6)$$

To generate $\alpha\beta$ stationary vector frame from $d-q$ rotating vector Park's transformation matrix is used given in Eq. (7).

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \cos(\theta) & \sin(\theta) \end{bmatrix} \begin{bmatrix} d \\ q \end{bmatrix} \quad (7)$$

The phase angle θ required for the above transformation is found from grid voltage transformation as shown in Fig. 6b [10, 11].

The active component of current of I_d is compared with reference current I_d^* the error is fed to PI controller which make error zero so that required active power is fed to the grid. During this time, current I_q^* is made zero as reactive power control is not in the focus of the paper.

4 Filter Requirements

The cut off frequency of the filter used at the output side of the grid connected inverter should be 10 time of the grid frequency and 1.5 times of converter switching frequency. For an LCL filter to be used, the value of inductor on inverter side filter is given by the equation [9] as shown in (Fig. 7).

$$L_{\text{inv}} = \frac{V_{\text{dc}}}{16 \times f_s \times \Delta I_{L-\text{max}}} \quad (8)$$

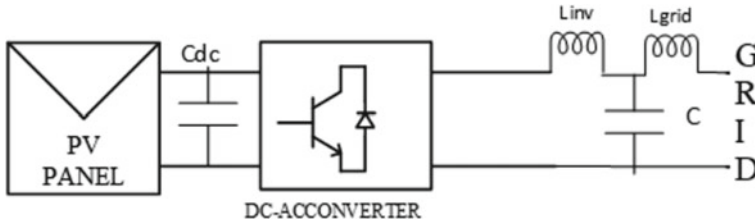


Fig. 7 Grid connected system indicating filter

V_{dc} is the dc link voltage at the input side of the inverter, f_s is switching frequency $\Delta I_{L-\max}$ ripple current that should be 10%.

$$L_{\text{grid}} = r \times L_i \quad (9)$$

L_{grid} is grid side inductance and L_i inverter side inductance where r is the factor of proportionality usually 0.6.

$$C = 0.05 \times C_b \quad (10)$$

where

$$C_b = \frac{1}{\omega n \times Z_b} \quad (11)$$

Z_b is base impedance and C is the filter capacitance.

Filter values, L_{grid} , L_{inv} are grid inductor, inverter inductor and C is the capacitance value calculated from above equations, values found as $L_g = 4$ mH, $L_{\text{inv}} = 4$ mH and $C = 6.6$ μF .

5 Simulation Results and Analysis

The output of the of the inverter, has three level voltage, the ripple current is very minimum, and the frequency of the inverter is same as switching frequency. Inverter produces very small leakage current as the common mode voltage is found to be zero. The simulation is carried for the parameters shown in Table 2.

Figure 8 indicates common mode voltage, proving that its value for HERIC grid connected inverter is constant whose value is 250 V in this case for a DC voltage of 500 V. The leakage current depends on the variation of common mode voltage across parasitic capacitor which shown as constant so that leakage current is zero as shown in Fig. 9.

Table 2 Simulation parameters [3]

Simulation parameter	Value
DC input voltage	500 V
Grid Voltage	230 V
Grid frequency	50 Hz
Switching frequency	2.4 kHz
Filter Grid inductor L_{grid}	4 mH
Filter inverter inductor L_{inv}	4 mH
Filter Capacitor	6.6 μ F
Parasitic capacitance	0.1 μ F
Active power rating	3000 W

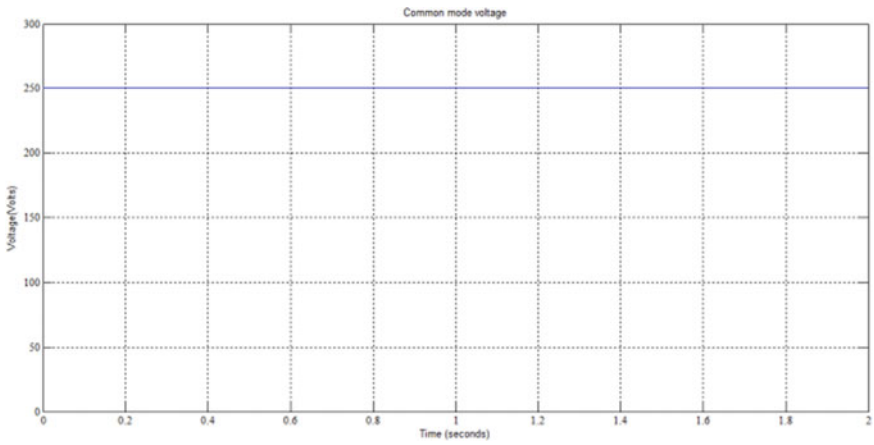


Fig. 8 Common mode voltage

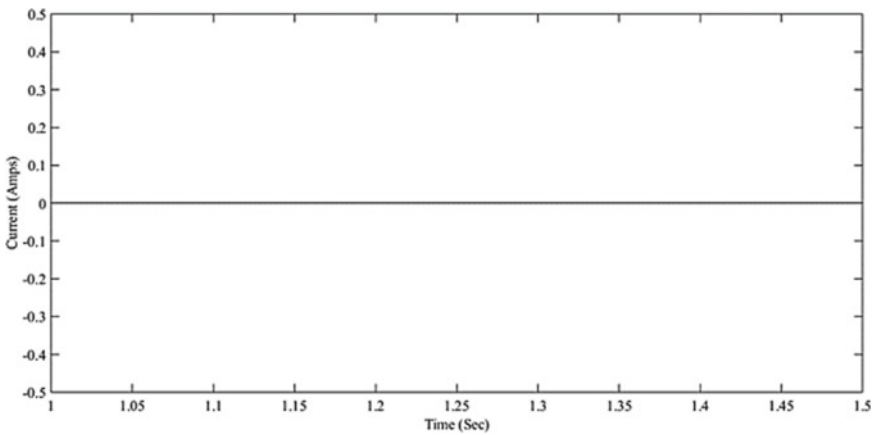


Fig. 9 Leakage current in HERIC grid connected inverter

Figure 12 shows the grid voltage along and current that is fed from inverter indicating voltage and current are in phase. The current profile is shown under steady state condition and it is in phase with voltage wave form.

The output voltage of HERIC inverter shown in Fig. 10 with a dc peak value of 500 V and a switching frequency of 2.4 kHz. Figure 11 indicates the inverter output current with the application of LCL filter shown in Fig. 12.

Active power fed to the grid from PV system after a maximum power extracted from the system, shown in Fig. 13 with a value of 3000 W. Simulation is carried for a Dc voltage of 500 V, resulting a voltage equals to grid voltage so that synchronization between grid and inverter. The switching frequency of the inverter used as 2.4 kHz, a passive LCL filter is used to gain reduced harmonics in the system.

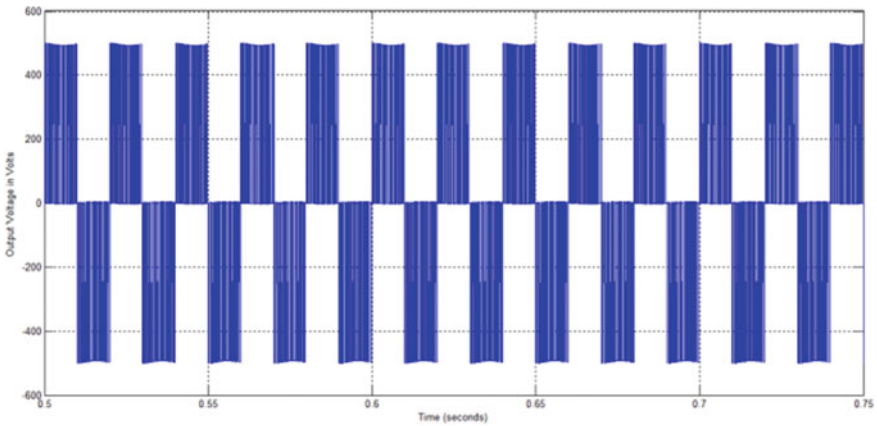


Fig. 10 Inverter output voltage

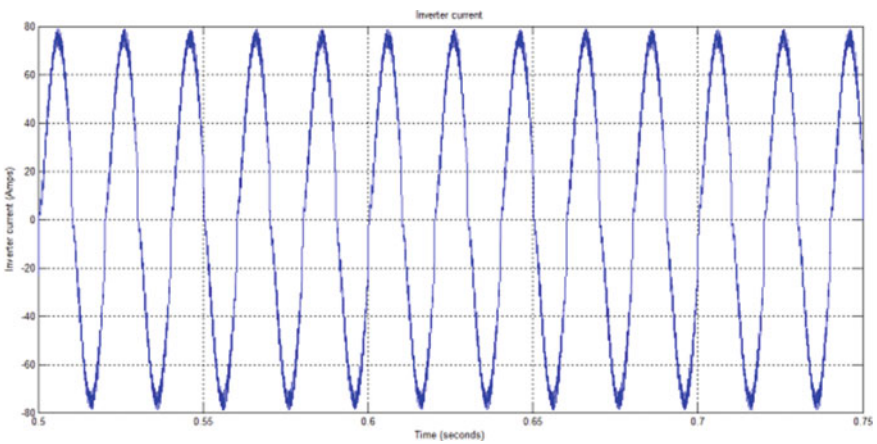


Fig. 11 Inverter output current

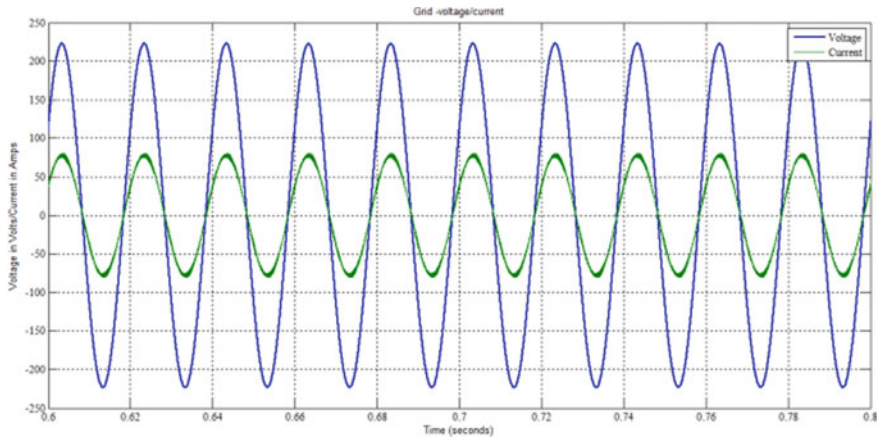


Fig. 12 Grid voltage-current

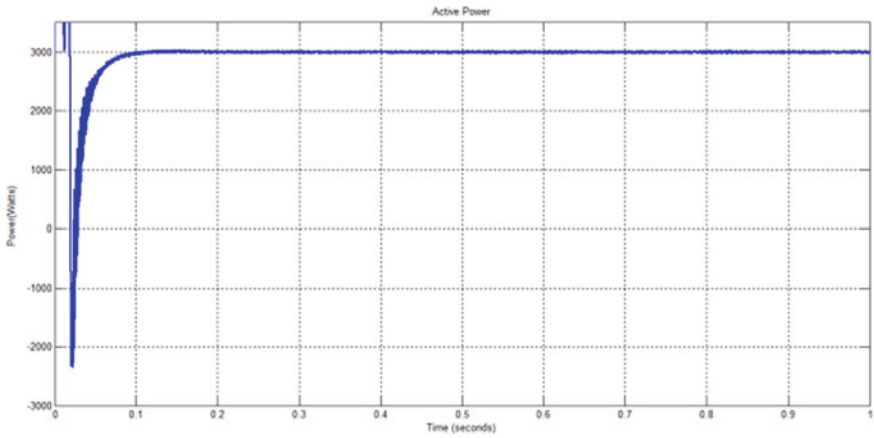


Fig. 13 Active power supplied by PV

6 Conclusion

Transformerless inverters will give increased efficiency which when compared with the inverters that consisting transformer. HERIC inverter is found to be one of the most effective converters for grid connection as the common mode voltage is constant for this topology and thus leakage current is minimum making a low electromagnetic interference. A simulation is carried to find the leakage current and to inject the active current form PV to grid.

References

1. R.A. Mastromauro, M. Liserre, T. Kerekes, A.D. Aquila, A single-phase voltage-controlled grid-connected conditioner functionality. *IEEE Trans. Ind. Electron.* **56**(11), 4436–4444 (2009)
2. T.F. Wu, H.S. Nien, C.L. Shen, T.M. Chen, A single-phase inverter system for PV power injection and active power filtering with nonlinear inductor consideration. *IEEE Trans. Ind. Appl.* **41**(4), 1075–1083 (2005). <https://doi.org/10.1109/TIA.2005.851035>
3. M.N.H. Khan, M. Forouzes, Y.P. Siwakoti, L. Li, T. Kerekes, F. Blaabjerg, Transformerless inverter topologies for single-phase photovoltaic systems: a comparative review. *IEEE J. Emerg. Sel. Top. Power Electron.* (2020). <https://doi.org/10.1109/JESTPE.2019.2908672>
4. E. Akpinar, A. Balikci, E. Durbaba, B.T. Azizoglu, Single-phase transformerless photovoltaic inverter with suppressing resonance in improved h6. *IEEE Trans. Power Electron.* **34**(9), 8304–8316 (2019). <https://doi.org/10.1109/TPEL.2018.2886054>
5. M. Premkumar, T.R. Sumithira, Design and implementation of new topology for solar PV based transformerless forward microinverter. *J. Electr. Eng. Technol.* **14**(1). <https://doi.org/10.1007/s42835-018-00036-2>
6. G. Janardhan, G.N. Srinivas, N.N.V. Surendra Babu, Realization of constant common mode voltage in transformerless photo voltaic inverter topologies, in *2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET)*, pp. 1–5 (2018). <https://doi.org/10.1109/ICCSDET.2018.8821224>
7. T. Kerekes, R. Teodorescu, P. Rodríguez, G. Vázquez, E. Aldabas, A New high-efficiency single-phase transformerless PV inverter topology. *IEEE Trans. Ind. Electron.* **58**(1), 184–191 (2011)
8. N. Yanshu, Single phase sine wave PWM inverter circuit simulation and the design of filter based on Matlab. *MATEC Web Conf.* **232**, 4–7 (2018). <https://doi.org/10.1051/mateconf/201823204030>
9. A. Kahlane, L. Hassaine, M. Kherchi, LCL filter design for photovoltaic grid connected systems, pp. 227–232 (2014)
10. M.R. Bengourina, M. Rahli, S. Slami, L. Hassaine, PSO based direct power control for a multifunctional grid connected photovoltaic system. *Int. J. Power Electron. Drive Syst.* **9**(2), 610–621 (2018). <https://doi.org/10.11591/ijpeds.v9n2.pp610-621>
11. S. Paghdar, U. Sipai, K. Ambasana, P.J. Chauhan, Active and reactive power control of grid connected distributed generation system, in *Proceedings of the 2017 2nd IEEE International Conference on Electrical, Computer and Communication Technologies. ICECCT 2017*, no. February 2017 (2017). <https://doi.org/10.1109/ICECCT.2017.8118019>
12. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol. 661, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
13. I.P. Antchev, Mihail, PII for single phase grid connected inverters. *Int. J. Electr. Eng. Technol.* **4**(5), 56–77 (2013)
14. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON)* held at Deenbandhu Chhotu Ram University of Science and Technology, SONEPAT, India on Feb 28–March 1, 2020
15. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
16. G. Janardhan, N.N.V. SurendraBabu, Transformerless Photo voltaic inverter topologies for low power domestic applications. *CVR J. Sci. Technol.* **11** (2016)
17. M.E. Ropp, S. Gonzalez, Development of a MATLAB/simulink model of a single-phase grid-connected photovoltaic system. *IEEE Trans. Energy Convers.* **24**(1), 195–202 (2009)

18. Y. Tang, W. Yao, P.C. Loh, F. Blaabjerg, Highly reliable transformerless photovoltaic inverters with leakage current and pulsating power elimination. *IEEE Trans. Ind. Electron.* **63**(2), 1016–1026 (2016)
19. R.N. Shaw, P. Walde, A. Ghosh, A new model to enhance the power and performances of 4x4 PV arrays with puzzle shade dispersion. *Int. J. Innov. Technol. Explor. Eng.* **8**(12) (2019)
20. M. Islam, S. Mekhilef, H6-type transformerless single-phase inverter for grid-tied photovoltaic system. *IET Power Electron.* (2015). <https://doi.org/10.1049/iet-pel.2014.0251>

Data Modeling for Energy Forecasting Using Machine Learning



Monika Goyal  and Mrinal Pandey

Abstract The consumption of energy in enormous amounts is one of the prime areas of concern for the researchers globally. The focus is on buildings as they are the biggest consumers of energy. The analysis of energy consumption pattern of buildings diverts our attention toward the heating, ventilation, and air conditioning system which according to studies accounts for maximum energy consumption within buildings. This research employs several regression techniques to predict the energy consumption of HVAC plants. The algorithms from various families of machine learning like linear models, function-based learning models, lazy learning techniques, and tree-based techniques were applied on the dataset obtained from a commercial building and the results were evaluated and compared based on performance metrics—root mean square error, mean square error, mean absolute error, and coefficient of determination. The interpretation of results obtained shows that K-nearest neighbor model from lazy learning family outperforms other regression models with reduced errors and improved accuracy.

Keywords Energy forecasting · Machine learning · Regression

1 Introduction

The consumption of energy within buildings demands focused analysis since several studies worldwide prove buildings to be the largest consumers of energy. Buildings are accountable for nearly 40% of the whole energy consumed in different sectors [1–4]. Apart from buildings, industry and transport account for major energy consumption. To handle the prevalent problem of energy wastage and to conserve energy for the generations to come, the consumption of energy in buildings must be analyzed. Research on energy consumption patterns within buildings reveals that

M. Goyal (✉) · M. Pandey
Department of Computer Science and Technology, Manav Rachna University, Faridabad, India

M. Pandey
e-mail: mrinal@mru.edu.in

the largest portion of energy, i.e., 40–50% is consumed by heating, ventilation, and air conditioning (HVAC) system [5, 6]. This motivates the researchers to focus on energy consumption by the HVAC system.

The purpose of HVAC system is to maintain the required ambient temperature in a building. Several components contained in HVAC system are chillers, cooling towers, primary pumps, and secondary pumps. Each component consumes energy to operate and perform its designated task. All HVAC components collaborate to maintain temperature and control relative humidity, thereby providing comfort to the building occupants. This increases the complexity of the problem as energy consumption is to be optimized without compromising on occupant's comfort.

Machine learning has been quite popular and effective in dealing with the issue of energy consumption in buildings. Regression techniques are efficient in predicting the energy consumed by HVAC and by buildings. There are many regression techniques with different features and they can predict results with reduced errors and improved accuracy. Although regression models have been applied in previous researches for energy prediction in buildings, very few models have been selected in each study. This research aims at applying various regression models on a real dataset of the cooling tower of a commercial building and analyzes the obtained results. Various parameters of cooling tower, i.e., dry bulb temperature, wet bulb temperature, relative humidity, cooling tower inlet temperature, and cooling tower outlet temperature were fed as input to the regression models, and the models predicted the power consumed by the cooling tower. The results obtained from each of the models were evaluated using certain well-known performance metrics- RMSE, MSE, MAE, and R-squared. The objective of this study is to explore various well-known regression models and analyze their performance by applying all of them to the same dataset.

Further in this paper, literature review is presented in Sect. 2 and overview of machine learning techniques used in this research are presented in Sect. 3, followed by research approach in Sect. 4. Section 5 describes the experiments performed and results obtained followed by Conclusion in Sect. 6.

2 Literature Review

Numerous studies can be found in literature where the researchers have applied machine learning techniques for predicting the consumption of energy in buildings and efforts toward optimizing its use. Machine learning techniques can be broadly classified into unsupervised techniques and supervised techniques. This section describes the application of various techniques from these two categories. Table 1 explores the work done by applying unsupervised machine learning techniques in the area of energy consumption and optimization in buildings. The application and results obtained after applying supervised machine learning techniques are described in Table 2 [7, 8].

Table 1 Publications relevant to unsupervised methods

Publication	ML technique/algorithm	Data source	Important features	Performance	Results
D'Oca and Hong [7]	Clustering (K-means) followed by association (FP growth)	An office building in Frankfurt am main, Germany	Window opening and closing	Not discussed	Occupants in manually ventilated buildings largely influence the indoor environment and energy performance
Xiao and Fan [8]	Clustering (entropy-weighted K-means) association (Apriori)	A commercial building in Hong Kong	Power consumption of various components in a building	Time taken in computing each step lesser than 120 s	The framework is suitable for enhancing building operational performance
Li et al. [9]	Clustering (K-means, EWKM, PAM, fuzzy C-means), association (Apriori)	R410A 29.8 kW type VRF	Energy consumption and saving potential in HVAC systems	Optimal results of clustering obtained by k-means by keeping k at 3	Identified patterns of energy consumption and extracting rules of energy consumption in VRF systems
Carreira et al. [5]	Clustering (K-means)	Simulated environment using weather of Lisbon, Portugal	RFID system, occupant votes, HVAC management	Energy reduction of 26.79% than baseline	Automatic increase/decrease in temperature based on user votes

Table 2 Publications relevant to supervised methods

Publication	Algorithm	Data source	Important features	Performance	Results
Sha et al. [10]	MLR, ANN, SVM	A retail building in China	Degree day	RMSE with SVR less than 20%	ANN and SVM outperformed MLR for cooling model
Li et al. [11]	ANN	Simulated dataset of buildings	Method of characterization decomposition and method of spatial homogenization decomposition	Energy consumption deviation is within 10%	Early stage accurate prediction of energy consumption by complex buildings
Zhao et al. [12]	Classification (C4.5, Locally weighted Naïve Bayes, SVM) Regression (linear regression, locally weighted regression, SVR)	An office building in Pennsylvania, USA	Office appliance power consumption data	90.29%	Occupant behavior affects plug load energy consumption
Zeng et al. [13]	Classification (MLP neural network) Prediction (Boosting tree algorithm)	HVAC system operated by ERS in Ankeny, Iowa	Energy saving in HVAC	Accuracy of energy consumption model = 90% All temperature model = 99%	More energy saving if only temperature constraints considered ignoring humidity
Candanedo et al. [14]	MLR, SVM-radial, RF, GBM	Residential building in Belgium	Atmospheric pressure	RMSE = 66.65 MAPE = 38.29% using GBM	Predicts energy used by appliances
Deb et al. [15]	Classification (ANN) Regression (MLR)	Office buildings in Singapore	Energy use intensity	ANN more accurate with MAPE = 14.8 against MLR with MAPE = 19.11	Develops prediction models by selecting appropriate building variables

(continued)

Table 2 (continued)

Publication	Algorithm	Data source	Important features	Performance	Results
Gallagher et al. [16]	K-nearest neighbors, ANN, decision tree, MLR, SVM	Biomedical manufacturing facility	Energy consumed by compressors, cooling towers, water pumps in industry	51.09% reduction in error	Models constructed for measurement and verification of energy savings in industrial buildings
Jain et al. [1]	SVR	Residential building in New York	Temporal and spatial granularity	Coefficient of variation = 3.30%	Sensor-based energy forecasting of a residential building
Kusiak et al. [17]	CART, CHAID, BT, RF, MARSplines, MLP, MLP ensemble, SVM	IOWA Energy Center, Ankeny, IOWA	Supply air temperature, static pressure	7.66% total energy savings	Models built to optimize energy consumption
Manjarres et al. [18]	RF	Office building in Spain	Weather forecast, HVAC ON/OFF, mechanical ventilation	Energy consumption reduced by 48% for heating and 39% for cooling	Switches between HVAC and mechanical ventilation for energy savings
Marasco and Kontokosta [19]	Classification (falling rule list classifier)	Local Law 87 audit dataset, New York City including residential, office, hotel, warehouse buildings	Energy conservation measures	ROC AUC = 0.72–0.86	A model developed to provide a fast and economic alternative for energy audit in buildings
Park et al. [20]	Classification (decision tree)	Office buildings in South Korea	Gross floor area Building Use Ratio	All office buildings showed even improved operational ratings	Developed an energy benchmark to improve operational rating of office buildings

(continued)

Table 2 (continued)

Publication	Algorithm	Data source	Important features	Performance	Results
Rahman et al. [21]	DT, DA, SVM, KNN	Commercial building in Alexandria, VA	Compressor AHU	Accuracy of KNN 99% for combined compressor and AHU model	Technique for HVAC power disaggregation
Singaravel et al. [22]	RF, ERT, KNN, MLP	Weather data from Amsterdam, Brussels, Paris, London	Annual heating and cooling energy	Prediction errors lower than 8%	A component-based ML model developed for energy demand prediction
Tsanas and Xifara [23]	Linear regression Random forests	Simulated residential buildings assuming the environment of Athens, Greece	Heating load Cooling load	MAE = 0.51 for HL and 1.42 for CL	Framework developed to estimate HL and CL
Wei et al. [24]	CHAD, Exhaustive CHAD, BT, RF, MLP, MARSplines, SVM	Energy Resource Centre, Ankeny, Iowa	HVAC Indoor air quality	Average energy savings = 12.4% to 17.4%	An approach developed to optimize the total energy consumption of HVAC in office building

3 Machine Learning Techniques

Machine learning [25] is a concept that refers to making a machine learn, such that it behaves in a certain way when a particular type of input is fed and gives results accordingly. Classical machine learning techniques include the base models for prediction like multiple linear regression, support vector machines, K-nearest neighbors, artificial neural networks, etc. These techniques can be used to perform classification as well as regression. The algorithm families are described as follows:

3.1 Linear

In the linear regression model, a linear relationship exists between the variable to be predicted, known as the response variable or target variable, and one or more predictor variables that are fed as input to the algorithm. The linear model is expressed by the following Eq. (1).

$$E[Y] = \emptyset(X_1, X_2, \dots, X_K) \quad (1)$$

where Y is the response variable and X_1, X_2, \dots, X_K are predictor variables [26].

From the family of linear regression models, multiple linear regression has been chosen for experiments in this paper. Multiple linear regression is similar to simple linear regression, in which the response variable is dependent upon only one predictor variable. In MLR, predictor variables are more than one and processing is done on all the predictor variables to predict the response. The interpretation of results of MLR is comparatively complex, mainly due to correlation among predictor variables. For better predictions, the predictor variables should be chosen very carefully such that relevant ones are not missed and irrelevant ones are not used. The various indices which can be considered during MLR interpretation are regression weights, zero-order correlation coefficients, structure coefficients, relative weights, product measures, all possible subsets regression, dominance weights, and commonality coefficients [27].

3.2 Function Based

Also known as kernel-based learning models, these are robust and efficient. In these models, data points are embedded into a vector space known as feature space and relationships are derived among those points. The concept can be understood by Eq. (2)

$$\emptyset(x) = (\emptyset_1(x), \emptyset_2(x), \dots, \emptyset_N(x))^T \quad (2)$$

where $\emptyset(x)$ is called feature vector.

$\{x : x \in R^d\}$ is the feature space and $N \gg d$

$$\emptyset : R^d \rightarrow R^N \quad (3)$$

Equation (3) defines a function called feature map.

For a given feature map, associated kernel function k can be defined as in Eq. (4)

$$k(u, v) = \emptyset(u)^T \emptyset(v) \quad u, v \in R^d \quad (4)$$

where $k : R^d * R^d \rightarrow R$.

In kernel-based models, output depends on the selection of the appropriate kernel [28]. From this family of learning models, support vector regression has been chosen for this research. Four kernel functions related to SVR namely radial, linear, sigmoid and Polynomial have been implemented and compared.

3.3 *Lazy Learning*

Also known as “Memory-based Learning”, this method defers the processing of training dataset until a requirement to respond to a query arises. It employs a distance function to measure relevance and the nearby data points are considered more relevant [29]. K-nearest neighbor was chosen from the family of lazy learning to perform experiments for this research. It takes into consideration “k” number of closest instances in the training dataset to predict the value of the unknown instance. These k instances are found by applying a certain distance metric such that they are the k-nearest neighbors of the unknown instance [30]. The value returned is obtained after averaging the values of k-nearest neighbors [7, 31].

3.4 *Tree-Based Learning*

Tree-based models recursively partition a training sample over its predictor attribute space and fit a simple function to each of the resulting subgroups [32]. The algorithms selected from the tree family are random forests and classification and regression tree (CART).

Random forests are a tree-based technique where each tree depends on a set of random variables. For predicting a continuous variable using random forests, the trees are grown depending on θ , a random vector, in such a manner that $h(x, \theta)$ which is the tree predictor that takes on numeric values. The values of the response variable are numeric and it is assumed that the training sample is drawn independently from the distribution X of random vector Y [7, 33].

In CART, the regression tree is constructed by finding the heterogeneity existing in the dataset and then purifying the dataset. When the tree is grown, the split point and splitting attribute are determined by minimizing the mean squared error [32].

4 Research Approach

The energy consumed by HVAC indicates the overall energy consumed in buildings, because as mentioned earlier, HVAC is the largest energy consumer within a building, and buildings are the largest energy consumer worldwide. Therefore, the prediction of energy consumed in buildings is important and various ML techniques have been employed on historical records to forecast the future energy consumption. This study analyzes the records of various parameters related to cooling tower, which is a primary and important component of the HVAC plant and employs several techniques from various families in machine learning to predict the power consumed by cooling tower. The methodology followed in the research is shown in Fig. 1. Each phase of the methodology is explained in this section.

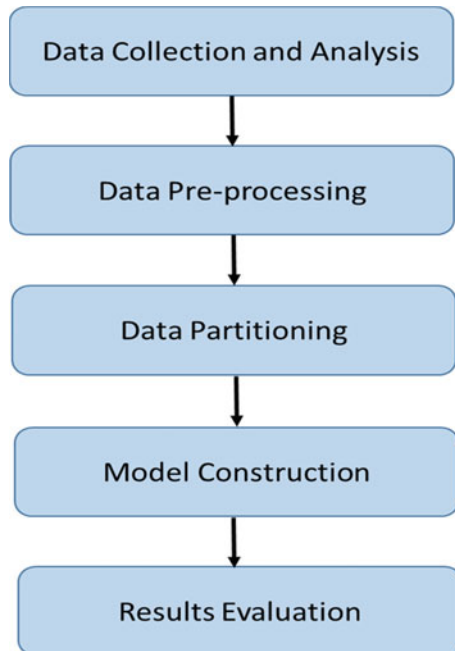


Fig. 1 Research approach

4.1 Dataset Collection and Analysis

The data of cooling tower collected from a hotel building in New Delhi is the focus of this study. It consists of samples which represent values of various parameters related to cooling tower. The values have been recorded by different sensors for different parameters at 5-min intervals for one year from Oct 2017 to September 2018.

Parameter Description. The obtained dataset has several parameters whose description can be found in Table 3 [25].

Out of the parameters described above, DBT, WBT, RH, CT_INLET, and CT_OUTLET have been taken as predictor variables for the regression algorithms and CT_POWER has been taken as response variable to be predicted.

Energy Consumption Patterns. The total energy consumed by cooling tower in each month is represented in Fig. 2.

After analyzing this energy pattern, the months can be categorized into two categories depending upon the amount of humidity present in the atmosphere. Months—May, June, July, August, September, October, and November can be considered as humid months, whereas the remaining months—December, January, February,

Table 3 Dataset parameter description

Parameter	Description	Unit
Dry bulb temperature (DBT)	Ambient temperature	°C
Relative humidity (RH)	Amount of water vapor in air relative to its temperature	%age
Wet bulb temperature (WBT)	Temperature brought down by water evaporation	°C
CT_INLET	Temperature of water entering into cooling tower	°C
CT_OUTLET	Temperature of water exiting from cooling tower	°C
CT_POWER	Power consumed by cooling tower	kW

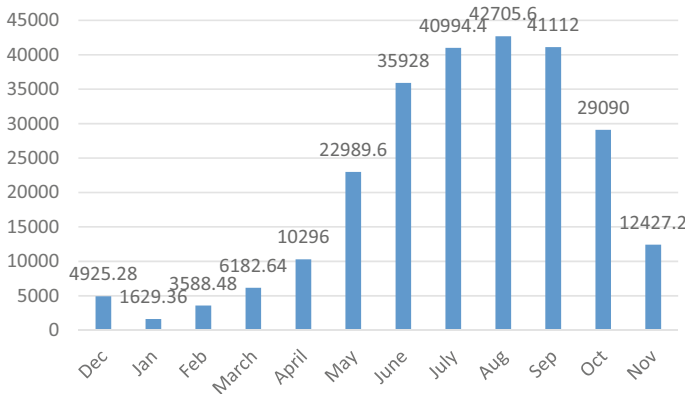


Fig. 2 Cooling tower energy consumption pattern for all months

March, and April are non-humid months. It is evident from the figure that energy consumption by cooling tower is more in humid months as compared to non-humid months because more power is consumed to control humidity. Based on these two weather conditions, i.e., humid and non-humid, one month was chosen from each category for further analysis. From the humid category, August was chosen as it is the most energy-consuming month while April was chosen from the non-humid category as it consumes maximum energy from its category. Figure 3 shows the daywise pattern of energy consumed by both the months—April and August [25].

It can be seen from the figure that energy consumption in August is consistently higher (around 1400 kW) for the whole month. For the month of April, energy consumption is more or less the same, ranging between 200 and 400 kW, for almost the whole month. A sudden spike can be seen toward the end of the month, because it is approaching toward May, when temperature and humidity start increasing.

Statistical Analysis. Statistical analysis was performed on the data of April and August and various statistics like minimum value, maximum value, mean, standard deviation, and variance for each input parameter are given in Tables 4 and 5, respectively.

The statistics given in the tables can be compared and certain facts can be deduced. The mean dry bulb temperature, i.e., the outside temperature is the same in April

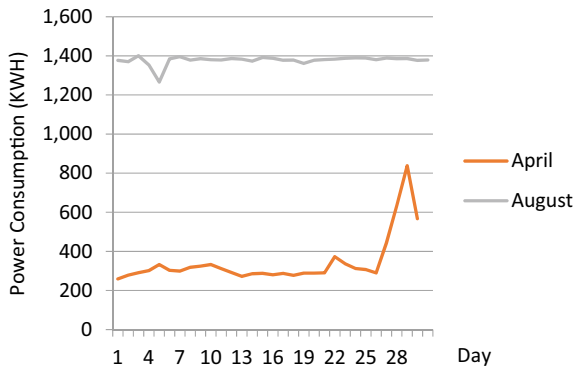


Fig. 3 Energy consumption pattern for April and August [34]

Table 4 Statistics for April

	DBT	RH	WBT	CT_INLET	CT_OUTLET
Min_Value	20.54	6.04	12.3	28.52	24.9
Max_Value	40.25	89.64	25.9	38.9	34.36
Mean	30.86	34.72	18.05	34.65	29.49
SD	4	14.56	2.4	1.13	1.25
Variance	16.08	212.17	5.77	1.28	1.57

Table 5 Statistics for August

	DBT	RH	WBT	CT_INLET	CT_OUTLET
Min_Value	24.03	47.71	23.84	30.86	27.6
Max_Value	36.42	98.77	31.87	39.52	36.42
Mean	30.37	73.92	27.57	36.6	32
SD	2.19	11.13	1.15	1.26	1.36
Variance	4.81	124	1.34	1.60	1.85

and August (around 30 °C). Despite this, there is a large difference in the energy consumption of both months. The reason behind this is the large difference in relative humidity that is present in the atmosphere. Average relative humidity is 34.72% in April, whereas it is 73.92 in August. Therefore, it is clear that most of the energy is consumed to combat humidity present in the air.

4.2 Preprocessing Phase

The quality of the constructed model for prediction depends on the records in the training dataset. Thus, the quality of the records is of utmost importance for increasing the accuracy of results and performance of the model. Generally, the obtained data is raw, scattered, unorganized, may contain missing values and outliers, is in an inappropriate format unsuitable for algorithm application. All these issues add complexity to the task of pattern extraction. Therefore, preprocessing of data is performed to deal with these issues and to simplify and improve the task of prediction. Though most of the readings in the dataset used in this research were at an interval of 5 min, still few extra readings were recorded at 1-min or 2-min intervals. These extra records were removed from the dataset to ensure symmetry while integrating different files containing values for different parameters. Few records had missing values for one parameter or the other, which were filled by the value from the previous record, because the value of the previous record was approximately the same as it was recorded just 5 min earlier. The obtained dataset had values for DBT, RH, CT_INLET, CT_OUTLET, and CT_POWER. A new parameter called WBT was derived from two existing parameters—DBT and RH because WBT is an important parameter on which cooling tower power consumption depends. The recorded data was in three different files—DBT and RH values were in one file, CT_INLET and CT_OUTLET were in another and CT_POWER values were in yet another file. Data from all these files was integrated into a single file and converted into CSV format. The type of data in the dataset was string, so it was transformed into numeric to make it suitable for applying regression algorithms.

4.3 Data Partitioning Phase

After preprocessing, the dataset was ready for further processing. Before applying ML algorithms, the dataset was partitioned into training set and testing set in the ratio of 70–30%, respectively. The dataset is partitioned such that training set can be used to train the machine by applying various techniques on it and the testing set is used to evaluate the accuracy of the developed model.

4.4 Model Construction

The model was constructed using various machine learning techniques and ensemble techniques for regression. Several families of regression models were explored by applying their algorithms on the dataset and results were derived. Since the data used in this research is continuous, regression models have been used for experiments. The objective of regression analysis is the construction of mathematical models for finding and describing relationships among variables. As mentioned in Sect. 3, the algorithms implemented in this research work are multiple linear regression, support vector regression, K-nearest neighbors, random forests, classification and regression trees, gradient boosting machines, extreme gradient boosting, and algorithm for bagging [7, 8, 25].

4.5 Evaluation Metrics

The applied machine learning models have been evaluated using four well-known evaluation metrics namely root mean square error, mean square error, mean absolute error, and R-squared [35, 36].

5 Experiments and Results

The above-mentioned regression algorithms were implemented on the dataset using R Studio and evaluated using various performance metrics—RMSE, MSE, MAE, and R-squared [8]. The results are shown in Table 6. Figures 4, 5, 6, 7 and 8 graphically represent the results obtained for MLR, SVR, KNN, RF, and CART, respectively.

In MLR, the values of error metrics RMSE and MAE are lower for August (3.44 and 0.95) as compared to April (6.14 and 4.23), whereas the value of coefficient of determination, i.e., R-squared is higher for April than for August (0.23 for April, 0.05 for August).

Figure 5 shows the results of SVR with Radial Kernel.

Table 6 Results of regression algorithms

Algorithm family	Algorithm name	RMSE		MSE		MAE		R-squared	
		April	August	April	August	April	August	April	August
Linear	Month								
	MLR	6.14	3.44	37.78	11.89	4.23	0.95	0.23	0.05
Functions	SVR (Kernel = Radial)	4.66	4.2	21.7	17.6	2.82	0.62	0.19	-47.85
	KNN ($k = 3$)	3.8	3.7	14.8	13.9	1.8	0.5	0.54	-1.65
Trees	RF	5.08	3.09	25.88	9.57	3.15	0.7	0.5	0.428
	CART	5.22	4.19	27.24	17.55	3.18	0.78	0.48	0.64

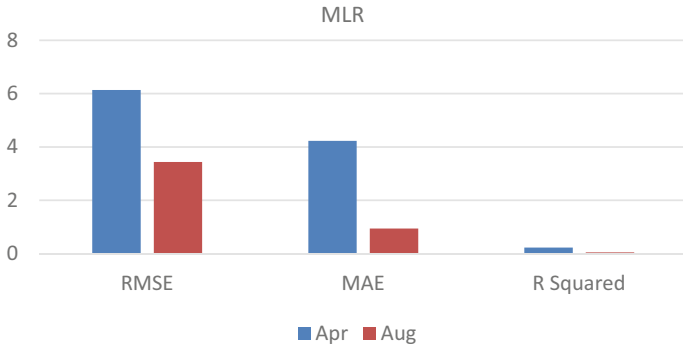


Fig. 4 Results of MLR

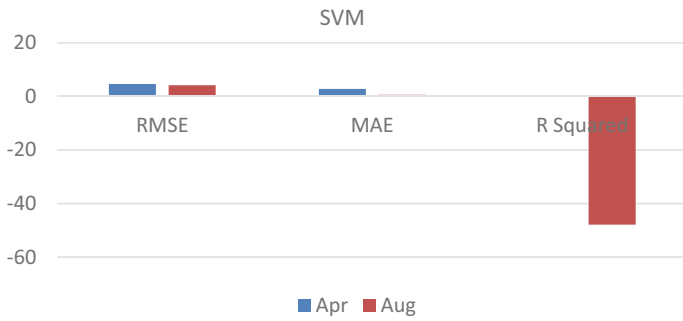


Fig. 5 Results of SVM

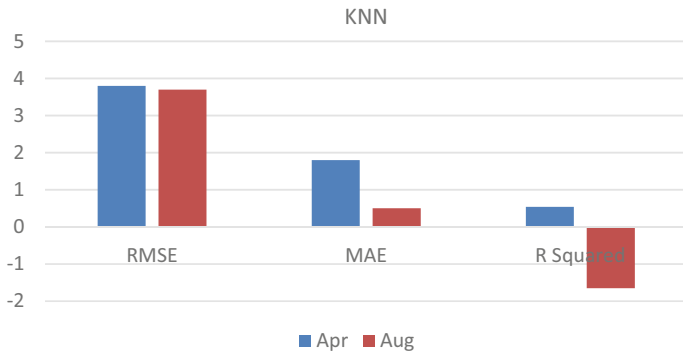


Fig. 6 Results of KNN

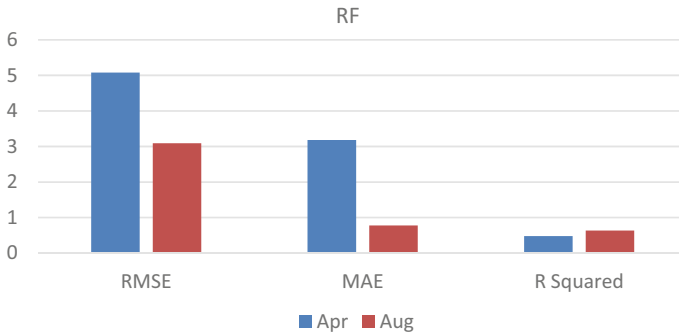


Fig. 7 Results of RF

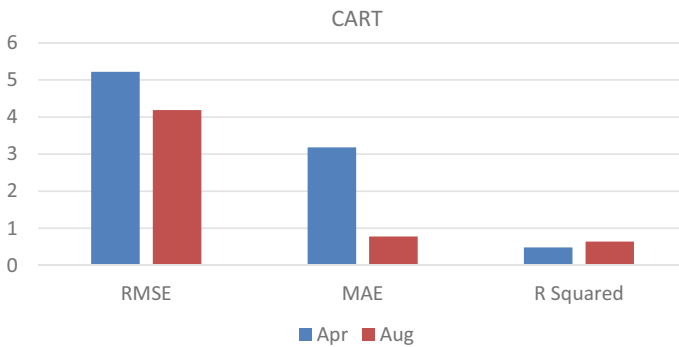


Fig. 8 Results of CART

The results of KNN with $k = 3$ are shown in Fig. 6.

The results of RF and CART are shown in Figs. 7 and 8. The results show that RF performs slightly better than CART when RMSE and MAE are considered. R-squared results for August have a higher value than RF.

6 Conclusion and Future Scope

This study works toward predicting the energy consumed by cooling tower, that is a key component in the heating, ventilation, and air conditioning system (HVAC). HVAC, being the largest energy consumer within a building, seeks attention for analysis to predict and consequently optimize the use of energy in buildings. Several regression models from various algorithm families have been implemented on a dataset obtained from a hotel building and compared based on certain evaluation metrics—RMSE, MSE, MAE, and R-squared. The results show that the K-nearest neighbor from the family of lazy learning techniques performs better with higher

accuracy. Future scope of this study will include ensemble Learning techniques for prediction.

References

1. R.K. Jain, K.M. Smith, P.J. Culligan, J.E. Taylor, Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy. *Appl. Energy* **123**, 168–178 (2014)
2. H. Naganathan, W.O. Chong, X. Chen, Building energy modeling (BEM) using clustering algorithms and semi-supervised machine learning approaches. *Autom. Constr.* **72**, 187–194 (2016)
3. M.W. Ahmad, M. Mourshed, Y. Rezugui, Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy Build.* **147**, 77–89 (2017)
4. J.S. Chou, D.K. Bui, Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. *Energy Build.* **82**, 437–446 (2014)
5. P. Carreira, A.A. Costa, V. Mansu, A. Arsénio, Can HVAC really learn from users? A simulation-based study on the effectiveness of voting for comfort and energy use optimisation. *Sustain. Cities Soc.* (2018)
6. J. Drgoña, D. Picard, M. Kvasnica, L. Helsen, Approximate model predictive building control via machine learning. *Appl. Energy* **218**, 199–216 (2018)
7. M. Goyal, M. Pandey, A systematic analysis for energy performance predictions in residential buildings using ensemble learning. *Arab. J. Sci. Eng.*, 1–14 (2020)
8. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
9. S. D'Oca, T. Hong, A data-mining approach to discover patterns of window opening and closing behavior in offices. *Build. Environ.* **82**, 726–739 (2014)
10. F. Xiao, C. Fan, Data mining in building automation system for improving building operational performance. *Energy Build.* **75**, 109–118 (2014)
11. G. Li, Y. Hu, H. Chen, H. Li, M. Hu, Y. Guo, M. Sun, Data partitioning and association mining for identifying VRF energy consumption patterns under various part loads and refrigerant charge conditions. *Appl. Energy* **185**, 846–861 (2017)
12. H. Sha, P. Xu, C. Hu, Z. Li, Y. Chen, Z. Chen, A simplified HVAC energy prediction method based on degree-day. *Sustain. Cities Soc.* **51**, 101698 (2019)
13. Z. Li, J. Dai, H. Chen, B. Lin, An ANN-based fast building energy consumption prediction method for complex architectural form at the early design stage, in *Building Simulation*, vol. 12, no. 4 (Tsinghua University Press, 2019), pp. 665–681
14. J. Zhao, B. Lasternas, K.P. Lam, R. Yun, V. Loftness, Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining. *Energy Build.* **82**, 341–355 (2014)
15. Y. Zeng, Z. Zhang, A. Kusiak, Predictive modeling and optimization of a multi-zone HVAC system with data mining and firefly algorithms. *Energy* **86**, 393–402 (2015)
16. L.M. Candanedo, V. Feldheim, D. Deramaix, Data driven prediction models of energy use of appliances in a low-energy house. *Energy Build.* **140**, 81–97 (2017)
17. C. Deb, S.E. Lee, M. Santamouris, Using artificial neural networks to assess HVAC related energy saving in retrofitted office buildings. *Sol. Energy* **163**, 32–44 (2018)
18. C.V. Gallagher, K. Bruton, K. Leahy, D.T. O'Sullivan, The suitability of machine learning to minimise uncertainty in the measurement and verification of energy savings. *Energy Build.* **158**, 647–655 (2018)

19. A. Kusiak, M. Li, F. Tang, Modeling and optimization of HVAC energy consumption. *Appl. Energy* **87**(10), 3092–3102 (2010)
20. D. Manjarres, A. Mera, E. Perea, A. Lejarazu, S. Gil-Lopez, An energy-efficient predictive control for HVAC systems applied to tertiary buildings based on regression techniques. *Energy Build.* **152**, 409–417 (2017)
21. D.E. Marasco, C.E. Kontokosta, Applications of machine learning methods to identifying and predicting building retrofit opportunities. *Energy Build.* **128**, 431–441 (2016)
22. H.S. Park, M. Lee, H. Kang, T. Hong, J. Jeong, Development of a new energy benchmark for improving the operational rating system of office buildings using various data-mining techniques. *Appl. Energy* **173**, 225–237 (2016)
23. I. Rahman, M. Kuzlu, S. Rahman, Power disaggregation of combined HVAC loads using supervised machine learning algorithms. *Energy Build.* **172**, 57–66 (2018)
24. S. Singaravel, P. Geyer, J. Suykens, Component-based machine learning modelling approach for design stage building energy prediction: weather conditions and size, in *Proceedings of the 15th IBPSA Conference*, pp. 2617–2626 (2017)
25. M. Goyal, M. Pandey, R. Thakur, Exploratory analysis of machine learning techniques to predict energy efficiency in buildings, in *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)* (IEEE, 2020), pp. 1033–1037
26. G.A. Seber, A.J. Lee, *Linear Regression Analysis*, vol. 329 (Wiley, 2012)
27. K.F. Nimon, F.L. Oswald, Understanding the results of multiple linear regression: beyond standardized regression coefficients. *Organ. Res. Methods* **16**(4), 650–674 (2013)
28. L.A. Belanche Muñoz, Developments in kernel design, in *ESANN 2013 Proceedings: European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning: Bruges (Belgium)*, pp. 369–378 (2013)
29. D.W. Aha (ed.), *Lazy Learning* (Springer Science & Business Media, 2013)
30. P. Cunningham, S.J. Delany, k-Nearest neighbour classifiers. *Multiple Classifier Syst.* **34**(8), 1–17 (2017)
31. F. Martínez, M.P. Frías, M.D. Pérez, A.J. Rivera, A methodology for applying k-nearest neighbor to time series forecasting. *Artif. Intell. Rev.* **52**(3) (2019)
32. N. Lin, D. Noe, X. He, H. Phoam, Tree-based methods and their applications, in *Springer Handbook of Engineering Statistics* (Springer-Verlag, London, 2006), pp. 551–570
33. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol. 661, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
34. M. Goyal, M. Pandey, Extreme gradient boosting algorithm for energy optimization in buildings pertaining to HVAC plants. *Energy Web* (2020)
35. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON)* held at Deenbandhu Chhotu Ram University of Science and Technology, SONEPAT, India on Feb 28–March 1, 2020
36. M. Goyal, M. Pandey, Towards prediction of energy consumption of HVAC plants using machine learning, in *International Conference on Recent Developments in Science, Engineering and Technology* (Springer, Singapore, 2019), pp. 254–265

Adaptive Linear Feedback Energy-Based Backstepping and PID Control Strategy for PMSG Driven by a Grid-Connected Wind Turbine



Youcef Belkhier, Abdelyazid Achour, Rabindra Nath Shaw, Walid Sahraoui, and Ankush Ghosh

Abstract Wind conversion system-based permanent magnet synchronous generator (PMSG) controller design is still a challenging work due to the PMSG nonlinear operation conditions and external disturbances. This work proposes a passivity-based control (PBC) associated with backstepping technique which ensures asymptotic convergence to the maximum power extraction, and stability of the closed-loop system that allows the PMSG to operate at an optimal speed and robustness of the system dynamics. The studied conversion system is constituted by a wind turbine, PMSG and buck-to-buck converter with DC-link voltage connected to the grid. The proposed method is used to control the generator-side converter, while a proportional–integral (PI) controller is used in the grid-side, to transmit only the active power to the distribution network. The objectives are achieved, and the reactive power and DC voltage quickly track their set values. The effectiveness of the proposed strategy is illustrated by numerical simulation results under MATLAB/Simulink.

Keywords Linear feedback control · Wind turbine · PMSG · Backstepping controller · Passivity-based control (PBC)

Y. Belkhier (✉) · A. Achour

Laboratoire de Technologie Industrielle et de l'Information (LTII), Faculté de Technologie, Université de Bejaia, 06000 Bejaia, Algeria

e-mail: youcef.belkhier@univ-bejaia.dz

R. N. Shaw

Department of Electronics and Communication Engineering, Galgotias University, Greater Noida, India

e-mail: r.n.s@ieee.org

W. Sahraoui

Laboratoire de Maitrise des Energies Renouvelables (LMER), Faculté de Technologie, Université de Bejaia, 06000 Bejaia, Algeria

A. Ghosh

School of Engineering and Applied Sciences, The Neotia University, Kolkata, West Bengal, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

177

S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,

Lecture Notes in Electrical Engineering 756,

https://doi.org/10.1007/978-981-16-0749-3_13

1 Introduction

Wind power generation is the most rapidly growing form of renewable energy where permanent magnet synchronous generator is the most used due to its advantages such as high efficiency, rapid response and high-power density [1, 2]. However, this kind of configuration presents some numerous issues which complicate the controller design for the PMSG such as parametric uncertainties, varying operation condition, nonlinear dynamics and external disturbances [3]. Control strategies for the PMSG have been mainly designed by nonlinear control methods. Literature [4] proposes an adaptive backstepping controller. Predictive control strategies for PMSG are illustrated in [5]. Reference [6] proposes a sliding-mode control approach. Fuzzy control method is developed by reference [7]. In this context, in this paper, we are interested in the so-called passivity-based control (PBC) due to its advantages of enhancing robustness properties, and the cancellation of system nonlinearities is avoided and energy-based approach [8]. In the PBC design, the objective is to bring the system to a desired dynamic, leaving the closed-loop system nonlinear, without cancelling dynamics or introducing controller singularities. A large publication that investigates a PBC on the PMSG can be found in the literature. Reference [9] reported a PBC-PI control. In [10], a passivity-based feedback control has been developed. Further, literature [11] presented a passivity-based sliding controller. Meanwhile, an adaptive passivity-based nonlinear control applied to PMSG is developed by [12]. Moreover, a passivity-based control combined with fuzzy integral sliding-mode control is investigated in [13].

This paper contributes with the new PBC with a linear feedback associated to backstepping technique, applied to the PMSG-based variable wind speed turbine with grid-connection via back-to-back converter to bring the PMSG to work at an optimal point while ensuring stability, fast convergence of the conversion system and performance improvement. This work aims that DC voltage and the reactive power must be maintained at their respective value, whatever the disturbances related to the PMSG. In Sect. 2, the model of the wind turbine and the PMSG is developed. Section 3 deals with the design procedure of proposed controller. Section 4 deals with the PI controller applied to the grid-side. Numerical results are provided in Sect. 5. Section 6 deals with the conclusion.

2 Wind Power and PMSG Modelling

2.1 *WindTurbine Model*

The mathematical model of the wind power that can be transformed by the turbine is given by (see Fig. 1) [1, 14, 15]:

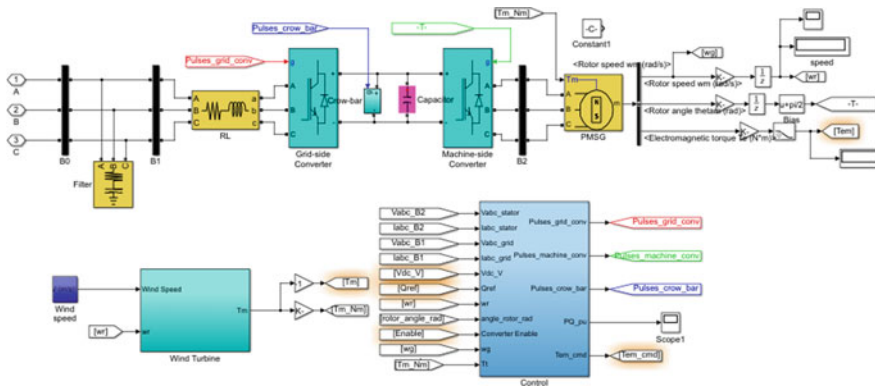


Fig. 1 Wind conversion system under Matlab/Simulink

$$P_m = \frac{1}{2} \rho C_p(\beta, \lambda) A v^3 \quad (1)$$

$$T_m = \frac{P_m}{\omega_m} \quad (2)$$

$$C_p(\beta, \lambda) = \frac{1}{2} \left(\frac{116}{\lambda_i} - 0.4\beta - 5 \right) e^{-\left(\frac{21}{\lambda_i}\right)} \quad (3)$$

$$\lambda_i^{-1} = (\lambda + 0.08\beta)^{-1} - 0.035(1 + \beta^3)^{-1} \quad (4)$$

where A represents the area of the blades, λ represents the tip-speed ratio, C_p denotes the coefficient of the power, β denotes the angle of the pitch, ρ denotes the wind density, v denotes the speed of the wind, ω_t represents the speed of the turbine and R denotes the blades radius.

2.2 PMSG Modelling

To design the proposed strategy, the dq model of the PMSG is considered and expressed as [16]:

$$v_{dq} = R_{dq} i_{dq} + L_{dq} \dot{i}_{dq} + p \omega_m \Im(L_{dq} i_{dq} + \psi_f) \quad (5)$$

$$J \dot{\omega}_m = T_m - T_e - f_{fv} \omega_m \quad (6)$$

$$T_e = \frac{3}{2} p \psi_{dq} \Im i_{dq} \quad (7)$$

where $R_{dq} = \begin{bmatrix} R_s & 0 \\ 0 & R_s \end{bmatrix}$ denotes the stator resistances matrix, f_{fv} represents the viscous friction coefficient, $\psi_f = \begin{bmatrix} \phi_f \\ 0 \end{bmatrix}$ represents the flux linkages, T_e represents the electromagnetic torque, $L_{dq} = \begin{bmatrix} L_d & 0 \\ 0 & L_q \end{bmatrix}$ denotes the stator inductions matrix, $v_{dq} = \begin{bmatrix} v_d & 0 \\ 0 & v_q \end{bmatrix}$ denotes the vector of the stator voltages, J denotes the inertia moment, $\mathfrak{S} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $i_{dq} = \begin{bmatrix} i_d \\ i_q \end{bmatrix}$ represents the vector of the stator currents.

3 Proposed Controller Design

Making the system passive is the principal aim of the proposed method. This is possible by introducing a damping term and reshaping its energy. The design of the adaptive linear feedback passivity-based backstepping controller (ALPBC) needs firstly to design the desired torque by the PID controller and then find voltage control as the control law of the PMSG by the backstepping controller.

3.1 Linear Feedback PBC Design

The relationship between the current vector i_{dq} and the flux linkage ψ_{dq} are given by [5]:

$$\psi_{dq} = \begin{bmatrix} \psi_d \\ \psi_q \end{bmatrix} = L_{dq} i_{dq} + \psi_f \quad (8)$$

Substituting the i_{dq} value obtained from (8) in (5) yields:

$$\dot{\psi}_{dq} + p\omega_m \mathfrak{S} \psi_{dq} = v_{dq} - R_{dq} i_{dq} \quad (9)$$

To design the proposed PBC, a linear feedback is used to the PMSG model that enables regulation in quadrature with the rotor flux and the armature reaction flux. This flux controller sets to zero i_d , and the PMSG model will be similar to that of a DC-generator. Then, the used controller has a linearizing feedback, which is expressed by:

$$v_d = -L_q p\omega_m i_q \quad (10)$$

Replacing (10) in (9) enforces the i_d current to zero that yields the simplified model expressed by:

$$\dot{\psi}_q - p\omega_m\psi_q = v_q - R_s i_q \quad (11)$$

In order to force current i_q to track its set value i_q^* , a proportional–integral (PI) is adopted, which yields:

$$v_q = k_{qp}(i_q^* - i_q) + k_{qi} \int (i_q^* - i_q) dt \quad (12)$$

where T_e^* represents the desired torque, and ψ_{dq} denotes the flux linkages expressed by [16]:

Where $k_{qp} > 0$ and $k_{qi} > 0$. To track the convergence of vector tracking error $\varepsilon_i = (i_q^* - i_q)$, the dynamic of the quadrature current is simplified, and its reference value is considered as the controller inputs of the PMSG when we assume that the PI current loop work satisfactorily by the proper choices of the positive gains. Then, the desired PMSG model, that is controlled by i_q^* decoupled by the application of the linear feedback, is expressed below:

$$\dot{\psi}_q - p\omega_m\psi_q = v_q - R_s i_q^* \quad (13)$$

$$J\dot{\omega}_m = T_m - T_e - f_{fv}\omega_m \quad (14)$$

$$T_e = \frac{3}{2}p\psi_q i_q^* \quad (15)$$

The desired model (13)–(15) is used to compute the controller input i_q^* with the proposed energy approach.

By taking the desired flux linkage vector as $\psi_{dq}^* = \begin{bmatrix} \psi_d^* \\ \psi_q^* \end{bmatrix}$, its tracking error vector as $e_f = \begin{bmatrix} e_{fd} \\ e_{fq} \end{bmatrix} = \psi_{dq} - \psi_{dq}^*$, and replacing e_f in (11), the expression of e_f is deduced. The Lyapunov theory is defined to compute the desired input i_q^* by using the defined function $V(e_f) = 0.5e_f^T e_f$. Then, it yields the following expression of i_q^* :

$$i_q^* = -R_s^{-1}(\dot{\psi}_q^* + p\omega_m\psi_q^*) + R_s^{-1}K_f e_{fq} \quad (16)$$

where $k_{fq} > 0$.

As can be seen, the current controller i_q^* depends on the desired quadrature flux ψ_{dq}^* . The PMSG operates at optimal torque only if i_d is fixed to zero. Under this condition, using (8), we get

$$\psi_q^* = L_q i_q^* \quad (17)$$

Then, from (7) and (17), it yields the following form of the desired quadrature flux ψ_{dq}^* :

$$\psi_q^* = \frac{2}{3} \frac{L_q}{p\varphi_f} T_e^* \quad (18)$$

The desired torque is computed by forcing the PMSG and wind turbine to rotate at the same speed and Eq. (7), which is given as follows:

$$T_e^* = J\dot{\omega}_m^* + T_m + f_{fv}\varepsilon_\omega \quad (19)$$

where ω_m^* represents the wind turbine speed. As can be seen from the above Eq. (19), it shows two drawbacks: the dependence of its convergence on the PMSG mechanical parameters (J, f_{fv}) and it is in open-loop. To solve that, a PID controller is proposed, to ensure fast convergence of the speed tracking error $\varepsilon_\omega = \omega_m^* - \omega_m$, eliminate the static error and introduce a robustness performance under different operation conditions:

$$T_e^* = J\dot{\omega}_m^* - k_{cp}\varepsilon_\omega - k_{ci} \int_0^t \varepsilon_\omega d\tau - k_{cd} \frac{d}{dt} \varepsilon_\omega \quad (20)$$

where $k_{cp} > 0$, $k_{ci} > 0$ and $k_{cd} > 0$.

3.2 Adaptive Backstepping Controller

The computation of control output of the PMSG is by applying the backstepping approach by stabilizing function generation. The linear feedback passivity-based strategy is used to generate the current controller i_{dq}^* expressed in (12), represents the current reference for the backstepping controller to generate the stabilizing function [4, 17]:

Firstly, the vector of current tracking error is defined as follow:

$$e_{dq} = \begin{bmatrix} e_d \\ e_q \end{bmatrix} = \begin{bmatrix} i_d^* - i_d \\ i_q^* - i_q \end{bmatrix} \quad (21)$$

Using (6), (12) and (21), the derivative of (21) is computed as below:

$$\dot{e}_{dq} = \begin{bmatrix} \dot{e}_d \\ \dot{e}_q \end{bmatrix}$$

$$= \left[\begin{array}{c} \frac{R_s}{L_d} i_d - \frac{p\omega_m L_q}{L_d} i_q - \frac{v_d}{L_d} \\ R_s^{-1} \frac{d}{dt} \left((-\dot{\psi}_q^* + p\omega_m \varphi_f) + K_f e_{fq} \right) \\ - \left(-\frac{R_s}{L_q} + \frac{p\omega_m L_q}{L_q} i_d - \frac{p\omega_m \varphi_f}{L_q} - \frac{v_q}{L_q} \right) \end{array} \right] \quad (22)$$

To design controller v_{dq} , a Lyapunov function is defined as follows:

$$V_e = \frac{1}{2} (e_d^2 - e_q^2) \quad (23)$$

The derivative of (23) yields:

$$\begin{aligned} V_e = & -k_1 e_d^2 - k_2 e_q^2 \\ & + e_d \left(\frac{R_s}{L_d} i_d - \frac{p\omega_m L_q}{L_d} i_q - \frac{v_d}{L_d} + k_1 e_d \right) \\ & + e_q \left(\begin{array}{c} \frac{R_s}{R_s + k_f L_q} \times \\ \left[\begin{array}{c} -R_s^{-1} \dot{\psi}_q^* + p\omega_m \varphi_f \left(\frac{f_{fv}}{J R_s} + \frac{1}{L_q} \right) \\ + i_q \left(\frac{R_s}{L_q} - \frac{(p\varphi_f)^2}{J R_s} \right) + \frac{p\varphi_f}{J R_s} T_m \\ - \frac{v_q}{L_q} + k_2 e_q \frac{R_s + k_f L_q}{R_s} \end{array} \right] \end{array} \right) \end{aligned} \quad (24)$$

After simplification, the Lyapunov function (24) becomes:

$$V_e = -k_1 e_d^2 - k_2 e_q^2 < 0 \quad (25)$$

where $k_{fd} > 0$ and $k_{fq} > 0$. The global asymptotic stability of the closed-loop is guaranteed by (25), and controller v_{dq} need to be chosen as below:

$$v_d = R_s i_d - p\omega_m L_q i_q + k_1 L_d e_d \quad (26)$$

$$\begin{aligned} v_q = & R_s^{-1} L_q \dot{\psi}_q^* + p\omega_m \varphi_f \left(\frac{L_q f_{fv}}{J R_s} + 1 \right) \\ & + i_q \left(R_s - \frac{L_q (p\varphi_f)^2}{J R_s} \right) + \frac{L_q p\varphi_f}{J R_s} T_m \\ & + k_2 L_q e_q \frac{R_s + k_f L_q}{R_s} \end{aligned} \quad (27)$$

Given the influence of the temperature on the stator resistance, and the PMSG dynamical, the stator resistance needs to be estimated adaptively. A new function is defined with the estimation error parameters as follows:

$$V_s = \frac{1}{2}(e_d^2 - e_q^2) + \frac{\Delta R_s^2}{2k_s} \quad (28)$$

where $\Delta R_s = R_s - \hat{R}_s$, \hat{R}_s represents the stator resistance estimated, and $k_s > 0$ represents the adaptive gain. Thus, the derivative of (28) gives:

$$\dot{V}_s = \dot{V}_e + \frac{\Delta R_s^2}{2k_s} \quad (29)$$

After simplification, Eq. (29) can be expressed as:

$$\dot{\hat{R}}_s = k_s \left(\frac{e_d i_d}{L_d} + \frac{e_q i_q}{L_q} \right) \quad (30)$$

4 Grid-Side Control

The mathematical model of the GSC is expressed below [18, 19]:

$$\begin{bmatrix} V_{id} \\ V_{iq} \end{bmatrix} = R_f \begin{bmatrix} i_{df} \\ i_{qf} \end{bmatrix} + \begin{bmatrix} L_f \dot{i}_{df} - \omega L_f i_{qf} \\ L_f \dot{i}_{qf} - \omega L_f i_{df} \end{bmatrix} + \begin{bmatrix} V_{gd} \\ V_{gq} \end{bmatrix} \quad (31)$$

$$C \dot{V}_{dc} = \frac{3}{2} \frac{v_{gd}}{V_{dc}} i_{df} + i_{dc} \quad (32)$$

where i_{df} and i_{qf} denote the grid currents, ω represents the grid angular frequency, V_{id} , V_{iq} denotes the inverter voltages, V_{gd} , V_{gq} denotes the grid voltages, L_f represents the filter inductance, R_f denotes the filter resistance, V_{dc} represents the DC-link voltage, C denotes the DC-link capacitor and i_{dc} represents DC current. The active power P_g and reactive power Q_g are given as:

$$\begin{cases} Q_g = \frac{3}{2} v_{gd} i_{qf} \\ P_g = \frac{3}{2} v_{gd} i_{df} \end{cases} \quad (33)$$

5 Simulation Results

Numerical validation is performed on the conversion by using Matlab/Simulink showed by Fig. 1. The closed-loop parameter values are listed in Table 1. The reference reactive power is fixed to zero, and the DC-link reference is set to 1150 V. Using the pole location method, the gains of the PBC-PI controller are $k_p = 100$, $k_i = 500$, $k_{qp} = 5$, $k_{qi} = 8000$. Gains of the DC-link are $k_{dcp} = 50$, $k_{dci} = 4000$, And current controller PI gains are $k_{cp} = 10$, $k_{ci} = 1000$ and $K_{fd} = K_{fq} = K_1 = K_2 = 100$ and $k_s = 70$.

The wind velocity is shown in Fig. 2. Figure 3 presents the pace of the electromagnetic torque response. Figure 4 shows the DC-link voltage, which quickly track its reference value 1150 V, without overshoot for the proposed strategy. Figures 5 and 6 show that the reactive power is very well kept at its zero-reference. The ALPBC

Table 1 System parameters

PMSG parameter	Value
Stator resistance (R_s)	0.006 Ω
Stator inductance (L_{dq})	0.3 mH
Wind turbine radius (R)	33.5 m
Wind density (ρ)	1.24 kg/m ²
Pole pairs number (p)	48
Flux linkage (φ_f)	1.48 Wb
Total inertia (J)	35,000 kg m ²
DC-link voltage (V_{dc})	1150 V
DC-link capacitor (C)	2.9 F
Grid-filter resistance (R_f)	0.3 pu
Grid-filter inductance (L_f)	0.3 pu
Grid voltage (V_g)	575 V

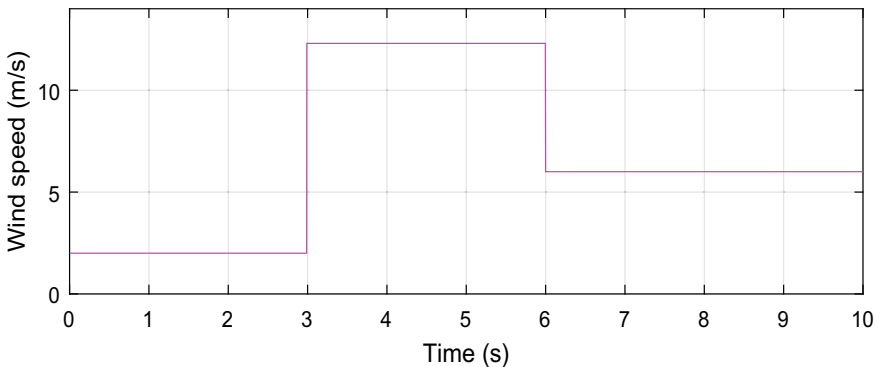


Fig. 2 Wind velocity

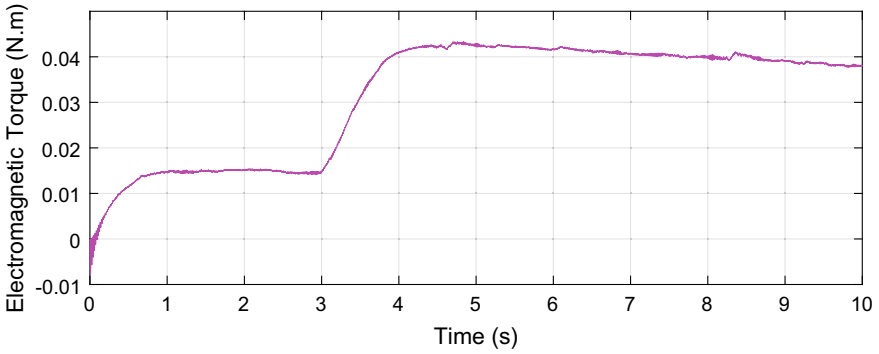


Fig. 3 Response of the electromagnetic torque

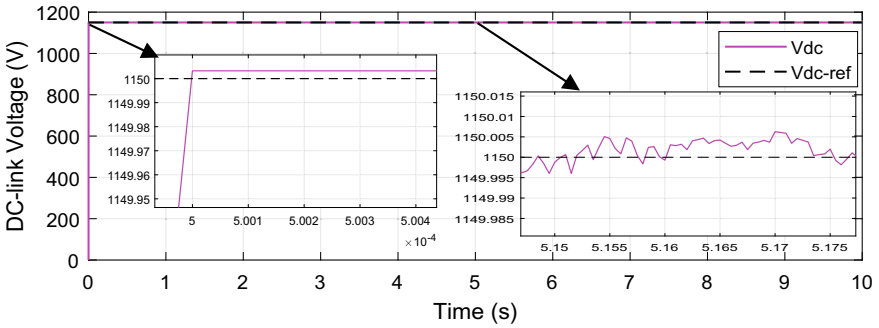


Fig. 4 Response of the DC-link voltage

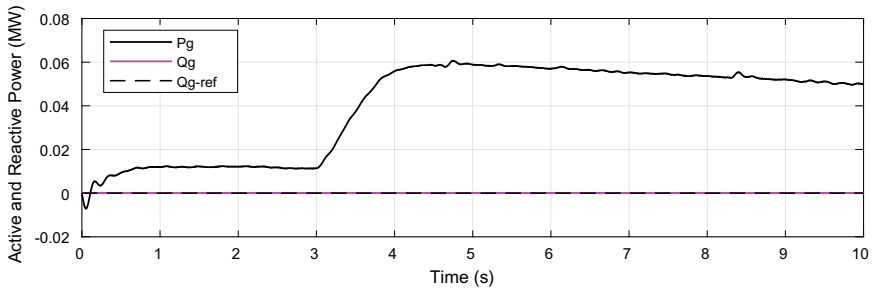


Fig. 5 Generated power

presents a fast convergence towards the references. The active power is only which is transmitted to the grid. In Fig. 7, we can see a perfect sinusoidal voltage absorption by the grid. This can be deduced that the control strategy successfully achieved main objectives of the paper. The ALPBC shows a higher torque, a reduced DC-link voltage error, a best power generation, with largely more power transmitted for the

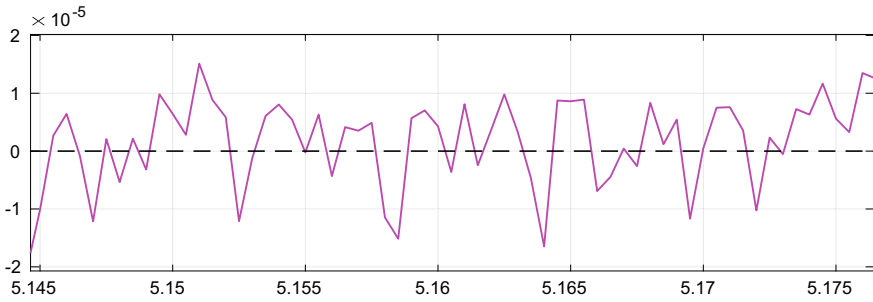


Fig. 6 Zoom on reactive power

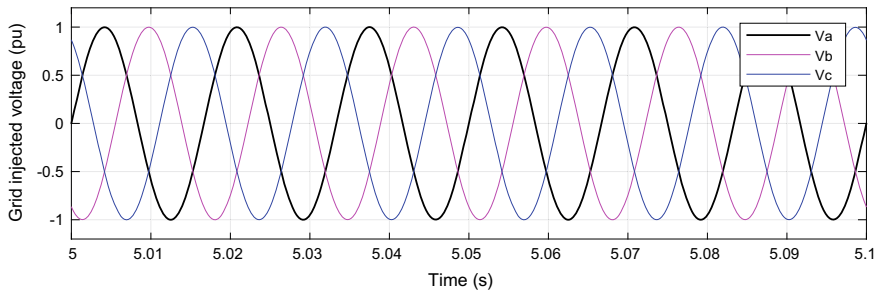


Fig. 7 Grid injected voltage

proposed one when compared to direct PI controller proposed by Sukati et al. in [2] and the passivity-based current control proposed by Belhier et al. in [20].

6 Conclusion

This paper contributes with an adaptive linear feedback energy-based backstepping controller for a wind power system with PMSG to improve its performances. The maximum power extraction from the wind turbine, by taking into account its entire dynamic when synthesizing the controller, represents the main motivation of the present work. The other aim of the study consists to maintain the generated reactive power and DC-link voltage at their reference values. The objectives of the paper are very well achieved. The conversion system shows high efficiency, high performance and fast convergence than the conventional strategies. The control strategy has simple structure and improves the performance of the conversion system.

References

1. Y. Belkhier, A.Y. Achour, R.N. Shaw, Fuzzy passivity-based voltage controller strategy of grid-connected PMSG-based wind renewable energy system, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, pp. 210–214, Greater Noida, India (2020). <https://doi.org/10.1109/ICCCA49541.2020.9250838>
2. K.S. Sukati, D.G. Dorrell, J. Agee, A direct PI controller without the feedforward terms for a VSC-based permanent magnet synchronous generator for a wind turbine, in *IECON 2018–44th Annual Conference of the IEEE Industrial Electronics Society*, pp. 5255–5260, Washington, DC, USA (2018)
3. X. Wang, S. Wang, Adaptive fuzzy robust control of PMSG with smooth inverse based dead-zone compensation. *Int. J. Control Autom. Syst.* **14**(2), 378–388 (2016)
4. J. Wang, D. Bo, X. Ma, Y. Zhang, Z. Li, Q. Miao, Adaptive back-stepping control for a permanent magnet synchronous generator wind energy conversion system. *Int. J. Hydrogen Energy* **44**(5), 3240–3249 (2019)
5. J.S. Lee, K.B. Lee, F. Blaabjerg, Predictive control with discrete Space-Vector modulation of Vienna rectifier for driving PMSG of wind turbine systems. *IEEE Trans. Power Electron.* **34**(12), 12368–12383 (2019)
6. S. Boulouma, H. Belmili, RBF neural network sliding mode control of a PMSG based wind energy conversion system, in *2016 International Renewable and Sustainable Energy Conference (IRSEC)*, IEEE, Marrakech, Morocco (2016)
7. Y. Krim, D. Abbes, S. Krim, M.F. Mimouni, Intelligent droop control and power management of active generator for ancillary services under grid instability using fuzzy logic technology. *Control Eng. Pract.* **81**, 215–230 (2018)
8. I.G. Polushin, D.J. Hill, A.L. Fradkov, Strict quasipassivity and ultimate boundedness for nonlinear control systems. *IFAC Proc.* **31**(17), 505–510 (2008)
9. R. Gao, I. Husain, R. Cisneros, R. Ortega, Passivity-based and standard PI controls application to wind energy conversion system, in *2016 IEEE Energy Conversion Congress and Exposition (ECCE)*, pp. 1–5, Milwaukee, WI, USA (2016)
10. B. Yang, T. Yu, H. Shu, D. Qiu, Y. Zhang, P. Cao, L. Jiang, Passivity-based linear feedback control of permanent magnetic synchronous generator-based wind energy conversion system: design and analysis. *IET Renew. Power Gener.* **12**(9), 981–991 (2018)
11. B. Yang, H. Yu, Y. Zhang, J. Chen, Y. Sang, L. Jing, Passivity-based sliding-mode control design for optimal power extraction of a PMSG based variable speed wind turbine. *Renew. Energy* **119**, 577–589 (2018)
12. J. Mash, M. Pahlevaninezhad, J. Praveen, Adaptive passivity-based nonlinear controller for wind energy conversion systems, in *2014 IEEE Applied Power Electronics Conference and Exposition-APEC*, pp. 1757–1764, Fort Worth, TX, USA (2014)
13. R. Subramaniam, Y.H. Joo, Passivity-based fuzzy ISMC for wind energy conversion systems with PMSG. *IEEE Trans. Syst. Man. Cyber. Syst.* to be published. <https://doi.org/10.1109/TSMC.2019.2930743>
14. G. Kapoor, A. Mishra, R.N. Shaw, S. Ajmera, Complex faults categorization technique for series-compensated transmission lines connected with wind-turbines based on DFWHT, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 703–708 (2020). <https://doi.org/10.1109/GUCON48875.2020.9231152>
15. M.D. Siddique, R.N. Shaw et al., Switched-capacitor based seven-level triple voltage gain boost inverter (7L-TVGBI), in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 848–852 (2020). <https://doi.org/10.1109/GUCON48875.2020.9231182>
16. A.Y. Achour, B. Mendil, S. Bacha, I. Munteanu, Passivity-based current controller design for a permanent-magnet synchronous motor. *ISA Trans.* **48**(3), 336–346 (2009)

17. A. Gaidi, H. Lehouche, S. Belkacemi, S. Tahraoui, M. Loucif, W. Guenounou, Adaptive backstepping control of wind turbine two mass model, in *Proceeding of the 6th International Conference on Systems and Control*, pp. 168–172, Batna, Algeria (2017)
18. Y. Belkhier, A.Y. Achour, Fuzzy passivity-based linear feedback current controller approach for PMSG-based tidal turbine. *Ocean Eng.* **218**, 108156 (2020). <https://doi.org/10.1016/j.oceaneng.2020.108156>
19. Y. Belkhier, A.Y. Achour, An intelligent passivity-based backstepping approach for optimal control for grid-connecting permanent magnet synchronous generator-based tidal conversion system. *Int. J. Energy Res.*, 1–16 (2020). <https://doi.org/10.1002/er.6171>
20. Y. Belkhier, A.Y. Achour, Passivity-based current control strategy for PMSG wind turbine, in *2019 IEEE 1st International Conference on Sustainable Renewable Energy Systems and Applications (ICSRESA)*, pp. 1–4, Tebessa, Algeria (2019). <https://doi.org/10.1109/ICSRESA49121.2019.9182518>

Permanent Magnet Synchronous Motor Drive for Electric Vehicle with Efficient Battery Management System



Varij Kumar and Saurabh Mishra

Abstract It is expected from permanent magnet synchronous motor (PMSM) to have rigidity with lower moment of inertia and power along with less mass. Because of its small size and higher power, PMSM has got proven as the one of the most favourable solutions for position control and speed controlling drives applied in machine tools and robots and also feasible for electric propulsion and control. Field-oriented control commonly called vector control is among the highly efficient control strategies of PMSM. This literature resembles the modelling of vector control of PMSM using an efficient battery management system in MATLAB/Simulink environment.

Keywords Vector control · DC–DC converter · Voltage source inverter · Switching pulse · State of charge

1 Introduction

DC motor drives prove quite easy and efficient control of speed as the field current and armature current are independent of each other [1]. Since AC drive system finds variety of applications such as electric vehicles, water pumping, fan-type loads and above 85% of aforementioned, induction motors are applied at present, but PMSM has the ability to replace them because of being more efficient, thereby reducing the replacement cost. Moreover, PMSM also possesses other speed control techniques but vector control got the attention for it being highly efficient [2–5]. Power electronics is applied at the initial stage itself where there is application of bidirectional DC–DC converter [6]. Electric vehicle finds a very bright future in itself, and the torque synchronization becomes the key aspect [7, 8]. In this paper,

V. Kumar · S. Mishra (✉)
Delhi Technological University, New Delhi 110042, India
e-mail: saurabhmishra@dtu.ac.in

V. Kumar
e-mail: varijkumar_2k17ee229@dtu.ac.in

there is a compilation of battery management system [9–14] of a higher efficiency with the field-oriented control of PMSM.

2 System Configuration and Machine Control Topology

A schematic control diagram is given in Fig. 1 comprising of battery for supply, bidirectional DC–DC converter to operate as per required DC link voltage, three-phase voltage source inverter (VSI) to provide AC supply to PMSM and gear box for appropriate shaft power transfer to the electric vehicle.

Speed PI controller is fed by the speed reference and speed feedback from speed sensor. Speed controller output is limited by the limiter to saturate q-axis current to maximum allowable current. Since flux weakening operation is not performed as electric vehicle is limited to a specific speed, d-axis current is kept zero. The main idea of employing vector control is to provide high efficiency speed control of PMSM. DC motors (shunt and separately excited) were the jack of industry before vector control advent for variable speed drives. To serve this purpose, there is requirement of controlling the gate pulses of fast switching inverter via controlling magnitude and phase of three stator currents (i_{as} , i_{bs} and i_{cs}) or in simple words current-controlled voltage source inverter is used here. Applying the inverted Park’s transformation, stator currents of PMSM (i_a^* , i_b^* , i_c^*) are derived from I_d^* and I_q^* .

$$i_a^* = I_q^* (\cos(\delta)) - I_d^* (\sin(\delta)) \tag{1}$$

$$i_b^* = I_q^* \left(\cos\left(\delta - \frac{2 * \pi}{3}\right) \right) - I_d^* \left(\sin\left(\delta - \frac{2 * \pi}{3}\right) \right) \tag{2}$$

$$i_c^* = I_q^* \left(\cos\left(\delta + \frac{2 * \pi}{3}\right) \right) - I_d^* \left(\sin\left(\delta + \frac{2 * \pi}{3}\right) \right) \tag{3}$$

$$\omega_{er} = \omega_{ref} - \omega_m \tag{4}$$

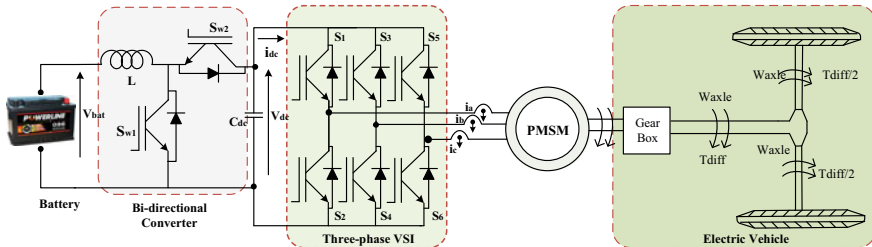


Fig. 1 Schematic diagram

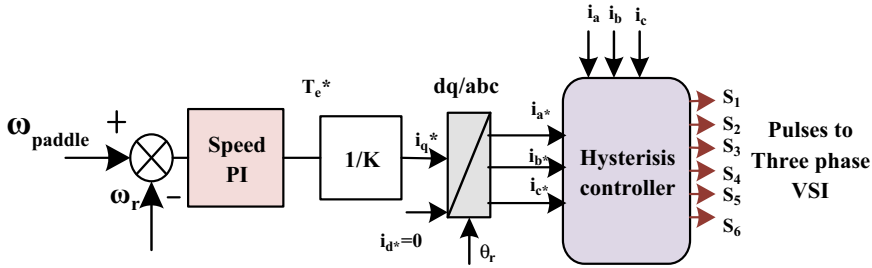


Fig. 2 Control schematic

$$T_e^*(n) = \left(k_p + \frac{k_i}{s} \right) \omega_{er}(n) \tag{5}$$

$$i_{q \text{ ref}}^* = \frac{T_e^* \text{ref}}{K} \tag{6}$$

$$K = \frac{3 * p * \lambda_{af}}{2} \tag{7}$$

where k_p is proportional gain, k_i is integral gain, ω_{er} is the error signal, p defines the number of pole pairs, and λ_{af} is flux linkage.

The reference currents hence obtained are fed to the current hysteresis controller along with the sensed motor phase currents to generate the switching gate pulse for voltage source inverter as per Fig. 2. The reference torque given to PMSM is kept constant although there is noticeable increase and dip while accelerating and de-accelerating the motor. The deceleration phase sees negative torque therefore negative current leading to charging of battery, thereby utilizing the braking power.

3 Battery Management System

The battery is coupled to the bidirectional DC–DC converter whose switching is controlled by the battery current controller output, i.e. appropriate duty cycle. Battery current controller is fed by two phases comprising of charging and discharging.

Discharging phase involves PI control as per the DC link voltage requirement. Charging phase involves separate PI control for charging at ample open-circuit voltage as per the state of charge (SOC). Given Fig. 4 shows the sum total of Fig. 3a, b.

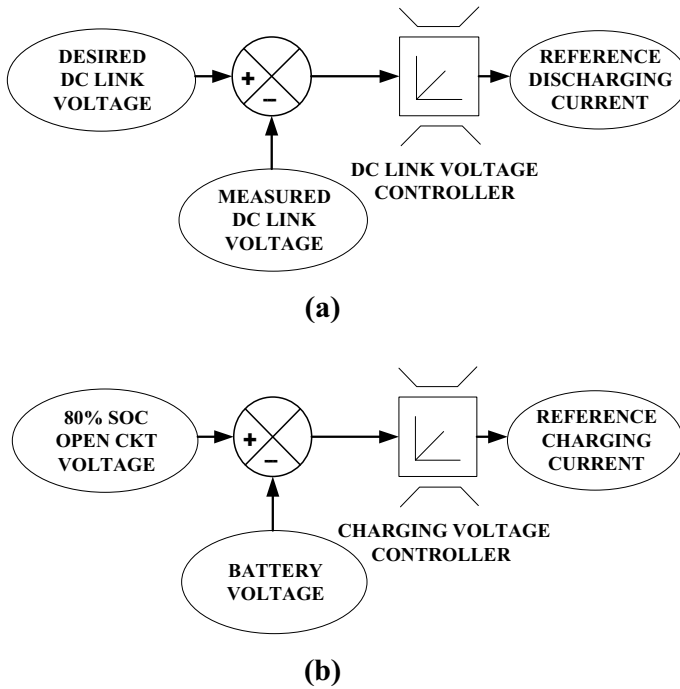


Fig. 3 a Discharging mode operation. b Charging mode operation

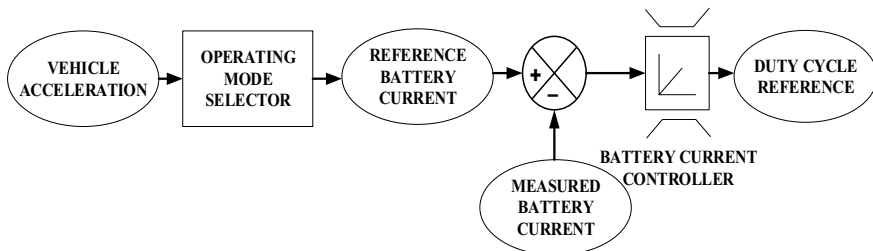


Fig. 4 Control schematic of DC–DC converter

4 Results and Discussion

The DC link voltage required is accurately achieved. Inverter and bidirectional DC–DC converter switching are also performing as per expectations. Rotor speed, electromagnetic and load torque are also in synchronism except while accelerating and de-accelerating. DC link or inverter input voltage variation with time as one can observe 400 V is rated, and it is achieved after practical overshoot of 6.25% at 0.75 s. Change in speed at 1 s is also causing a very small dip because of acceleration

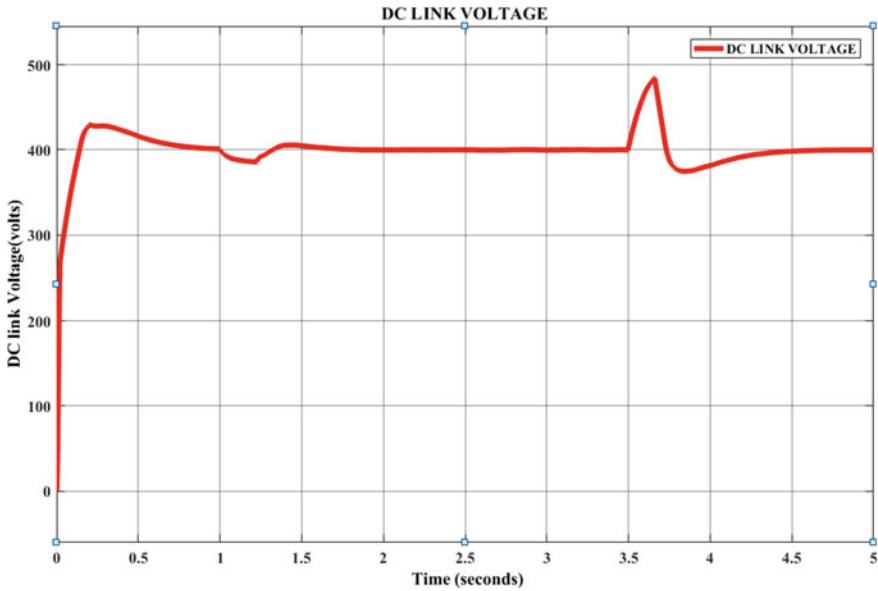


Fig. 5 DC link voltage

(Fig. 5) and at 3.5 s voltage shoots to 500 V as regenerative braking is observed with increment in SoC (Fig. 6) at 3.5 s (Table 1).

Battery power output (Fig. 7) is also stabilized at 450 W after 1.5 s, and current (Fig. 8) is drawn higher at acceleration but stabilization at 4 A is observed after peak of 12 A, and then at 3.5 s there is regenerative braking of PMSM and charging of battery is observed from 3.5 to 4 s as negative power and current is indicated.

Differential speed (Fig. 9) shows the switching mode of operation when positive it resembles acceleration of PMSM and discharging of battery while on negative it shows deceleration of PMSM and charging of battery. Similar kind of variation can be observed for q-axis stator currents (torque-producing component) (Fig. 10).

Load torque is always kept constant at 4 N m, electromagnetic torque (Fig. 11) shows significant increase at times of starting and acceleration, and there is negative torque fed back by PMSM while decelerating. Stator three-phase currents also show variation as per transient conditions (Fig. 12).

Reference speed is kept constant at 50 rad/s and at 1 s shifted to 110 rad/s and then descended to 40 rad/s. Measured speed traces quite practically. (Fig. 13).

Rotor angle is also showing its nature as per speed variation (Fig. 14).

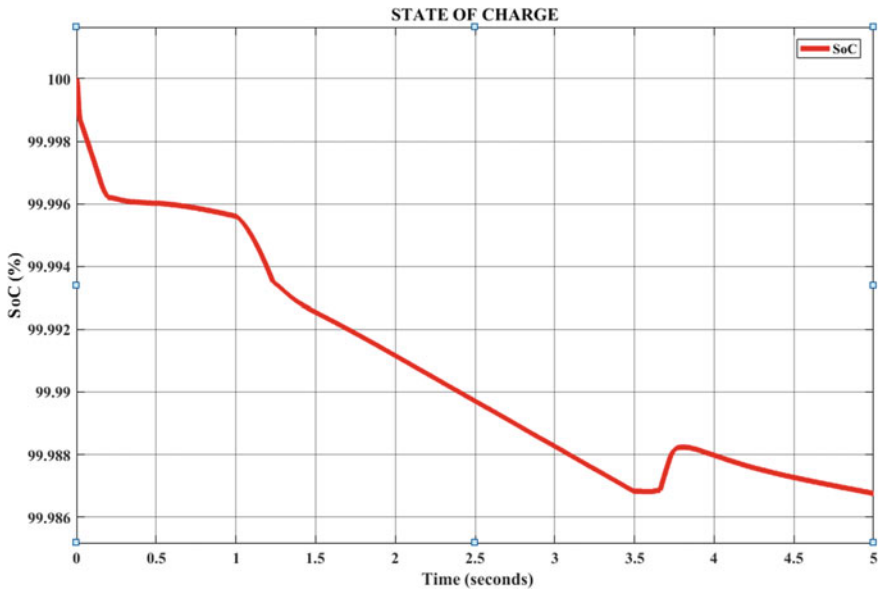


Fig. 6 State of charge of battery

Table 1 Discussion

Time (s)	Speed variation	Electromagnetic torque
0–1	Speed reference is kept constant at 50 rad/s However, the PMSM speed gradually increases from 0 to 50 rad/s in 0.16 s and remains constant	The reference torque was kept constant at 4 N m but the electromagnetic torque momentarily increases at the time of acceleration and stabilizes at 4 N m after 0.16 s Obviously, battery kept discharging to run the motor
1–3.5	Reference shoots to 110 rad/s while PMSM takes time to achieve the same and then stabilizes	Torque also faces an overshoot by 12 N m, thereby proving that acceleration requires higher torque and then stabilizes at 4 N m
3.5–5	Reference is allowed to descend to 40 rad/s PMSM enters braking mode	In braking mode, torque turns out to be negative and similarly the q-axis current, thereby charging the battery and enhancing the SOC

5 Simulated Parameters

See Tables 2, 3 and 4.

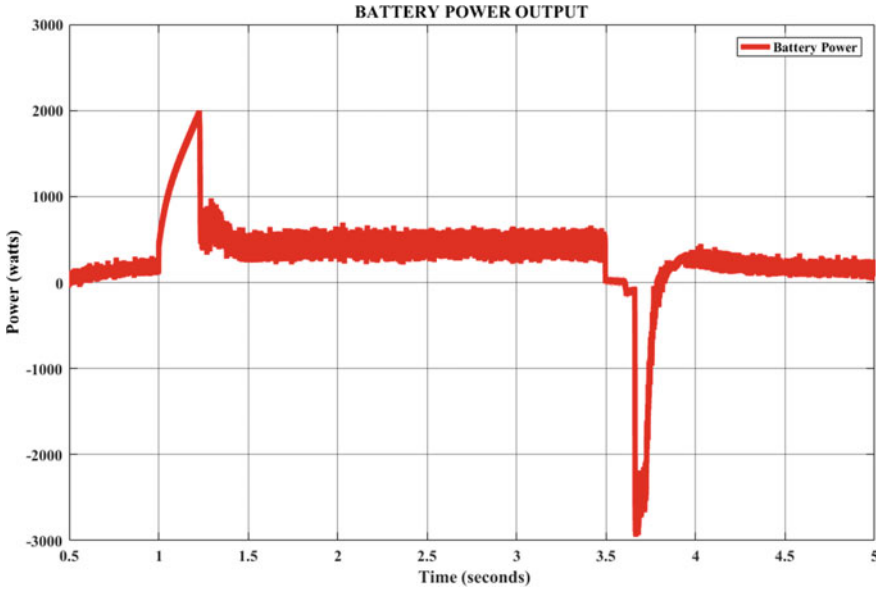


Fig. 7 Power output of battery

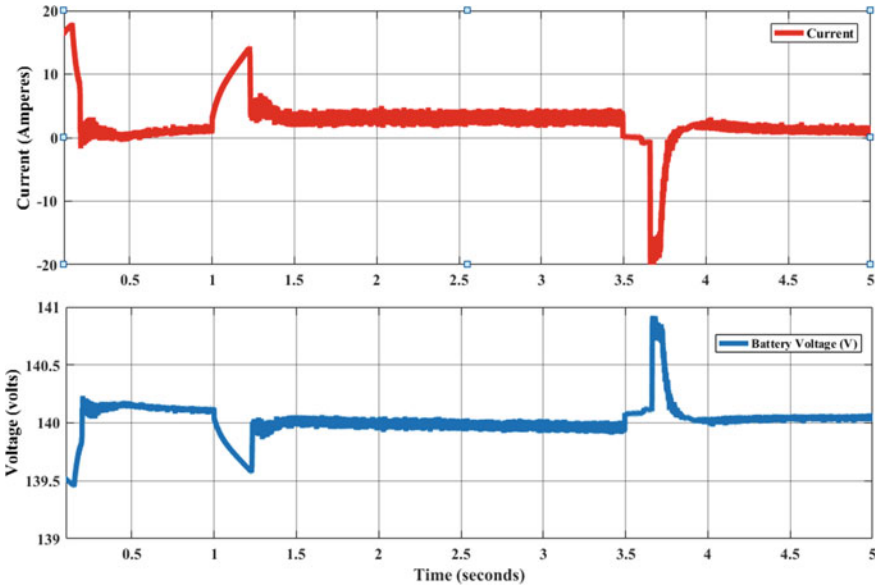


Fig. 8 Battery voltage and current

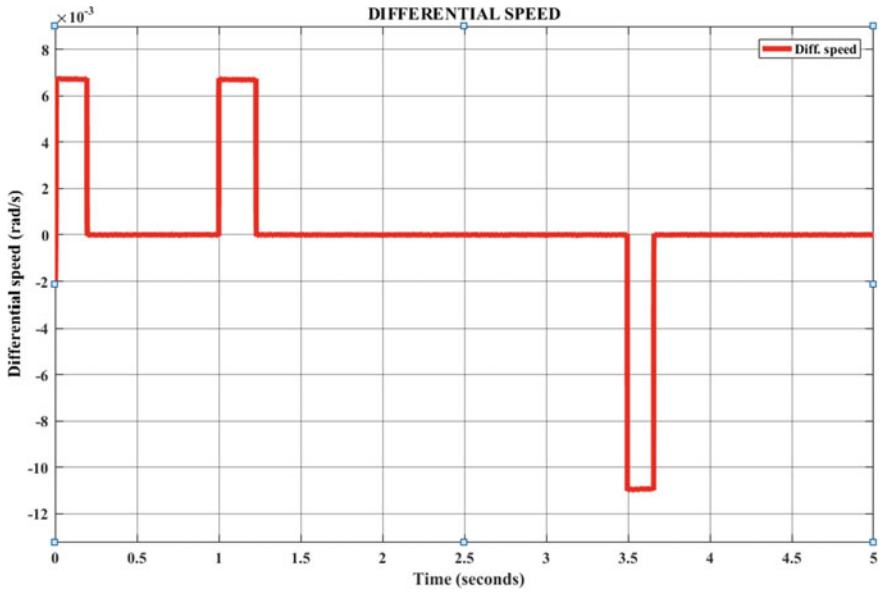


Fig. 9 Differential speed or mode selector of battery operation

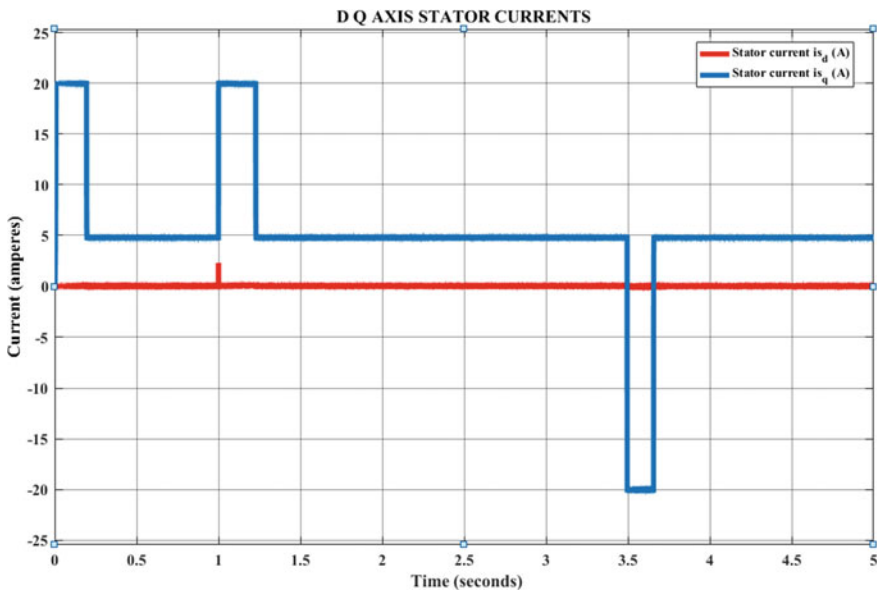


Fig. 10 D-Q-axis stator currents

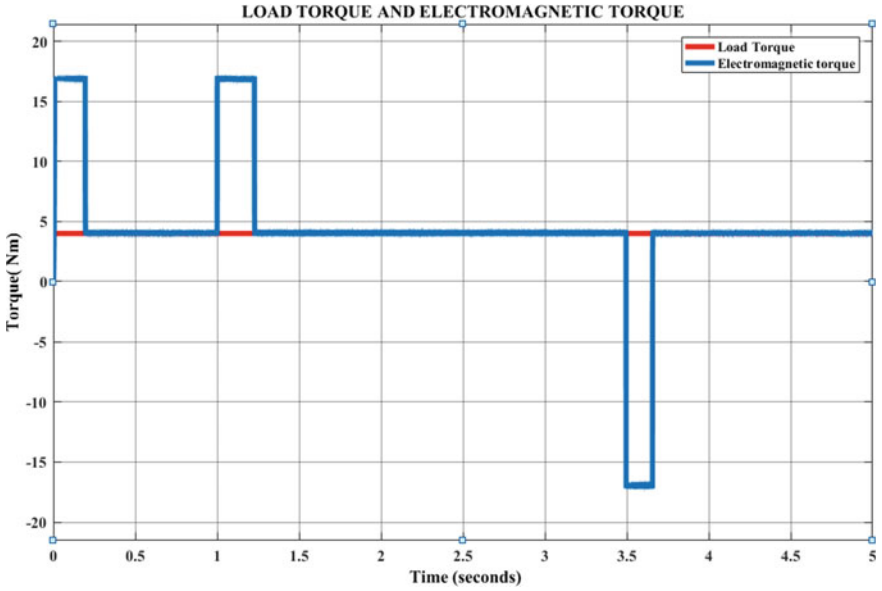


Fig. 11 Load and electromagnetic torque

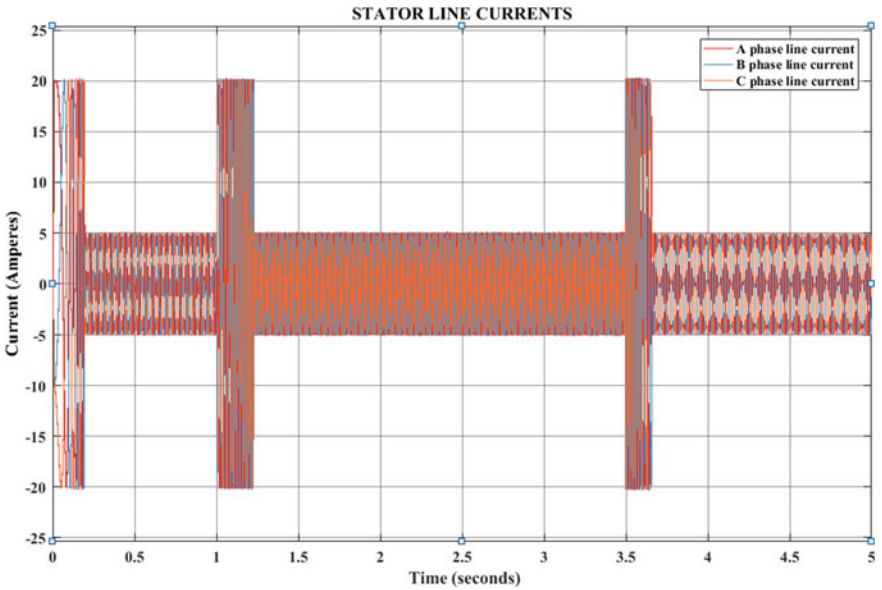


Fig. 12 Three-phase line current measurements

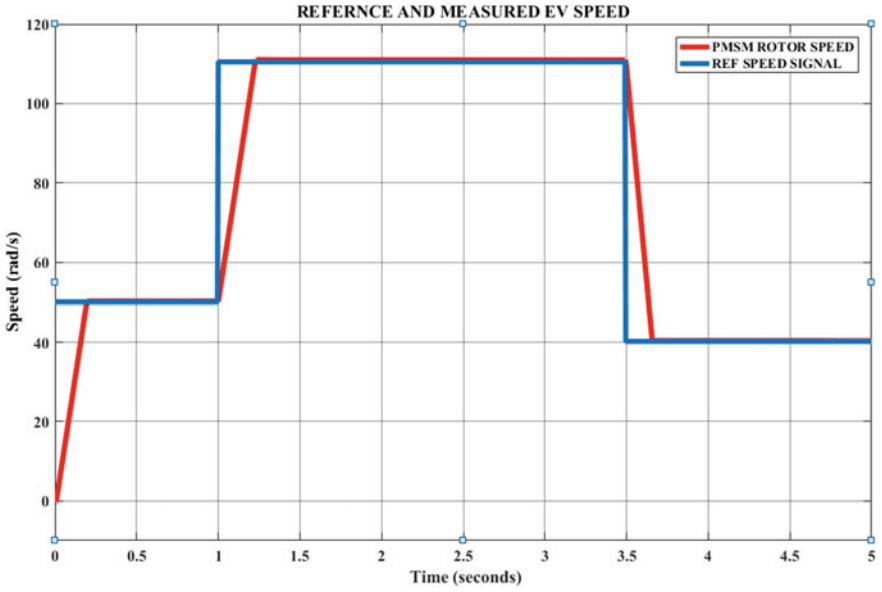


Fig. 13 Rotor and reference speed comparison

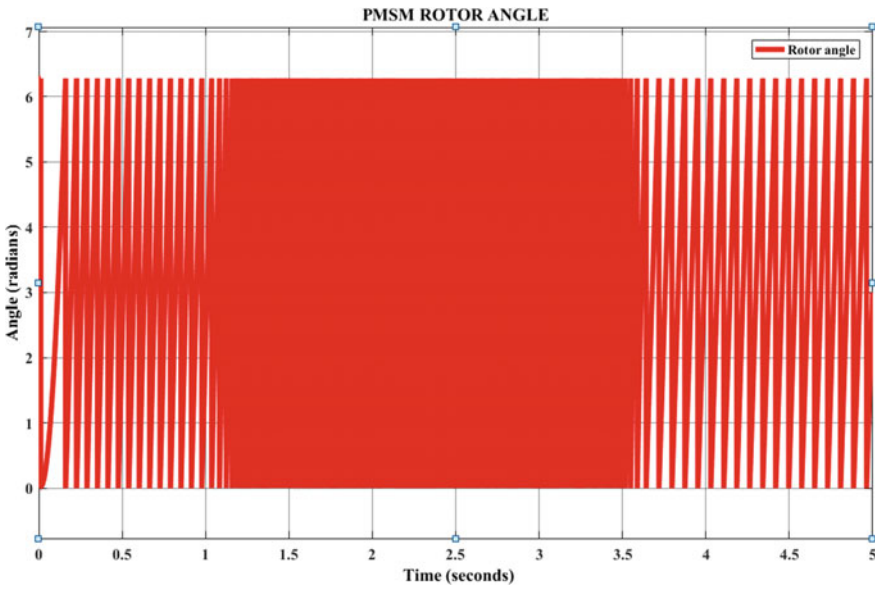


Fig. 14 Rotor angle

Table 2 Battery parameters

Quantity	Numerical Value
Fully charged voltage	120 V
Battery capacity	30 Ah
SoC	100%
Battery response time	30 s
Battery discharging PI controller	Proportional constant 0.4 Integral constant 2
Battery charging PI controller	Proportional constant 40 Integral constant 2000
Battery current PI controller	Proportional constant 0.1 Integral constant 0.55

Table 3 Bidirectional DC–DC converter

Quantity	Numerical Value
Series inductor	10 mH 0.05 Ω
Shunt capacitor	5 mF 1 mΩ
Switching frequency	20 kHz
Battery response time	30 s

Table 4 PMSM parameters

Quantity	Numerical Value
Rated power (P_{rated})	1.5 kW
Rated speed (ω_{rated})	157 rad/s
Poles	4
Direct axis inductance L_d	8.5 mH
Rated flux (φ_m)	0.141 Wb
Moment of inertia	0.048
Speed PI controller	Proportional constant 100 Integral constant 5
Saturation of output q-axis current	+20 A
Current hysteresis controller relay	+0.01

Acknowledgements We hereby acknowledge ICEEE Publisher, annexed sponsors and Delhi technological university for providing a good platform to showcase our findings.

References

1. V. Kumar, S. Mishra, Speed and current controller modelling for close loop control of separately excited DC motor for constant flux and flux weakening operation, in *IEEE Conference on Computing, Power and Communication Technologies (GUCON)*, Galgotias University, Greater Noida, UP, pp. 803–808, October 2020
2. P. Pillay, R. Krishnan, Development of digital models for a vector controlled permanent magnet synchronous motor drive, in *Conference Record of the 1988 IEEE Industry Applications Society Annual Meeting*, pp. 476–482, October 1988
3. P. Pillay, R. Krishnan, Modeling, simulation, and analysis of permanent-magnet motor drives. *IEEE Trans. Ind. Appl.* **25**, 265–273 (1989)
4. R. Krishnan, *Permanent Magnet Synchronous and Brushless DC Motor Drives*, 1st edn. (CRC Press, Boca Raton, FL, USA, 2010)
5. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
6. P. Freere, P. Pillay, Design and evaluation of current controllers for PMSM drives, in *IECON '90: 16th Annual Conference of IEEE Industrial Electronics Society*, pp. 1193–1198, March 1990
7. B. Bose, *Modern Power Electronics and AC Drives*, 2nd edn. (Pearson Education Pvt. Ltd., Singapore, 2001)
8. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol. 661, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
9. R.N. Shaw, P. Walde, A. Ghosh, A new model to enhance the power and performances of 4x4 PV arrays with puzzle shade dispersion. *Int. J. Innov. Technol. Explor. Eng.* **8**(12) (2019)
10. J. Larminie, J. Lowry, *Electric Vehicle Technology Explained*, 2nd edn. (Wiley, 2012)
11. R.N. Shaw, P. Walde, A. Ghosh, IOT Based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON)* held at Deenbandhu Chhotu Ram University of Science and Technology, SONEPAT, India on Feb 28–March 1, 2020
12. Y. Wang, Y. Liu, Electronic control system design and test of pure electric vehicle battery management system, in *2011 Second International Conference on Mechanic Automation and Control Engineering*, pp. 1289–1292, July 2011
13. F. Helling, S. Götz, T. Weyh, A battery modular multilevel management system (BM3) for electric vehicles and stationary energy storage systems, in *16th European Conference on Power Electronics and Applications (IEEE, 2014)*
14. D.D. Artakusuma, H. Afrisal, A.I. Cahyadi, O. Wahyunggoro, Battery management system via bus network for multi battery electric vehicle, in *International Conference on Electrical Engineering and Computer Science (ICEECS)* (IEEE, 2014), pp. 179–181

Two-Step Experimental Validation of Impact of Irregular Irradiance on Solar Photovoltaic System's Performance



Paresh S. Nasikkar , Chandrakant D. Bhos , and Javed K. Sayyad 

Abstract Photovoltaic (*PV*) system, when subjected to irregular irradiance caused by shading, can lead to noticeable power losses. In this paper, the impact of irregular solar irradiance on the *PV* system has been analyzed by operating the *PV* system under different levels of solar irradiance in MATLAB/Simulink. The results have been experimentally validated using a *PV* system with two *PV* panels of 75 W each in series. Further, the results are also validated using a *PV* emulator. The analysis shows a noticeable reduction of 55% in the maximum power extracted from the *PV* system due to irregular irradiance. It also results in a reduction of fill factor and efficiency of the system by 53% and 28%, respectively.

Keywords Photovoltaic · Solar · Shading · Maximum power · Fill factor

Abbreviations

I	Total current (A)
I_{pv}	<i>PV</i> current (A)
I_s	Saturation current (A)
I_d	Diode current (A)
q	Charge on an electron ($1.60217662 \times 10^{-19}$ C)
k	Boltzmann constant (1.380649×10^{-23} J/K)
V	<i>PV</i> cell voltage (V)
T	Temperature (K)

P. S. Nasikkar (✉) · C. D. Bhos · J. K. Sayyad
Department of Electronics and Telecommunication Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Lavale, Pune 412115, India
e-mail: paresh.nasikkar@sitpune.edu.in

C. D. Bhos
e-mail: chandrakant.bhos@sitpune.edu.in

J. K. Sayyad
e-mail: javed.sayyad@sitpune.edu.in

N_s	Number of cells connected in series
a	Constant of diode ideality factor
R_s	Series resistance (Ω)
R_p	Parallel resistance (Ω)
V_{oc}	Open-circuit voltage (A)
I_{sc}	Short-circuit current (A)
V_{MPP}	Voltage at maximum power point (V)
I_{MPP}	Current at maximum power point (A)
FF	Fill factor
η	Efficiency

1 Introduction

Immense interest in power generation is produced because of expansion in the populace and industries. Non-renewable energy sources like oil, gas, coal, atomic, and so forth are commonly utilized for power generation. Notwithstanding, endless issues like a discharge of ozone-depleting substances, natural issues, and so on are confronted with the utilization of petroleum derivatives. The sustainable power assets are these days broadly utilized for power generation all through the world to fulfill the immense need for power. This is expected the way that sustainable power sources do not influence the climate. Likewise, the cost of non-renewable energy sources is continually changing and these fills will drain after certain years. Regardless, such issues can be overcome by adopting sustainable energy sources, and these are uninhibitedly gotten from nature. Wind and solar energy of these environmentally friendly power assets are significant. These energies are broadly accessible in plentiful sum and these are liberated from the disadvantages which are unmistakable in non-renewable sources. Solar energy is generally a superior alternative in Indian situations, and consequently, it is broadly used in power applications because of its sustainability, cleanliness, low noise, and moderately simplicity of upkeep. *PV* cell straightforwardly transforms the solar energy over to an electrical one. Yet, it needs to address some difficulties like energy variance and low efficiency of the *PV* module as noted from several pieces of literature.

Many researchers have highlighted the issue of the impact of varying solar irradiance on the *PV* systems. The effect of partial shading and temperature on the *PV* system is simulated in PSPICE and experimentally validated by Joao Paulo et al. [1]. Belkaid et al. have worked on the MPPT algorithm for a partially shaded *PV* system and gained promising results [2]. Basaran et al. studied the effect of irradiance measurement sensors on *PV* systems and experimentally achieved the performance ratio of 83.69%, 82.48%, and 80.87% using reference cell, pyranometer, and PVGIS, respectively [3]. The hot spot effect caused due to partial shading was addressed by Dolara et al. [4]. Magare et al. analyzed the effect of seasonal irradiance variations on different *PV* technologies [5]. Many kinds of literature have addressed the issue

of partial shading and its impact on *PV* systems [6–14] Performance degradation due to shading was reported by Boukenoui [15]. Shaw et al. have studied effects of solar irradiance on load sharing of integrated *PV* systems and got promising results [16]. Kumar et al. did predictive data analysis based energy management [17]. Shaw et al. have also worked on shade dispersion concept [18]. Manjunath et al. reported multiple maxima points due to shading [19]. Shaw et al. implemented IOT based MPPT under shaded conditions and achieved significant improvements in maximum power [20]. Bana et al. studied the power generation from *PV* systems under uniform as well as non-uniform solar irradiance for different configurations of *PV* modules [21].

This paper is organized into six sections. Section 2 introduces mathematical modeling of the system. Section 3 presents the impact of irregular irradiance on the *PV* system using MATLAB/Simulink model. The experimental validations using hardware setup and emulator are discussed in Sects. 4 and 5, respectively. Finally, the concluding remarks are reported in Sect. 6.

2 Mathematical Modeling of System

P–N junction similar to a diode is formed in a *PV* cell. A *PV* cell is modeled as a current source and parallel diode (see Fig. 1). R_p and R_s imply obstruction to the electron flow from n region to p region and the current flow, respectively. For an ideal *PV* cell:

$$I = I_{pv} - I_d = I_{pv} - I_s \left(\exp \frac{qV}{kT} - 1 \right) \tag{1}$$

With consideration of R_p and R_s , the above equation can be expanded as

$$I = I_{pv} - I_s \left(\exp \frac{q(V + IR_s)}{N_s k T a} - 1 \right) - \frac{V + IR_s}{R_p} \tag{2}$$

MPP can be observed on nonlinear I – V and P – V characteristics (see Fig. 2).

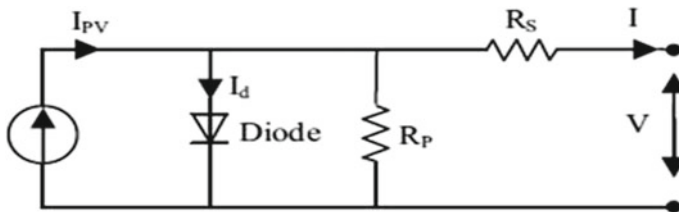


Fig. 1 Simplified model of *PV* cell

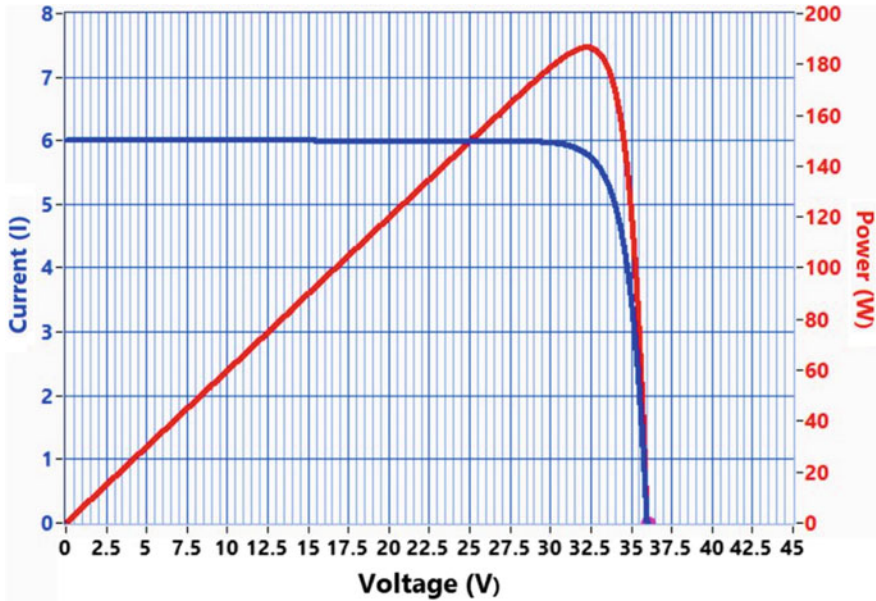


Fig. 2 I - V and P - V characteristics of a typical PV system

3 Impact of Irregular Irradiance on PV System

A MATLAB/Simulink model of a solar PV array has two series-connected PV panels (see Fig. 3). It is implemented for analyzing the impact of PS on PV panel performance. Parameters of said PV panel are tabulated in Table 1.

The simulated I - V and P - V curves with four different irradiance levels are illustrated (see Figs. 4, 5, 6, 7, 8, 9, 10 and 11) and the observations are tabulated in Table 2.

It can be observed from the results that irregular irradiance significantly reduces the maximum power extracted from the system.

4 Experimental Validation Using Two PV Modules

The PV system with two series-connected PV panels of 75 W each is used for experimental validation (see Figs. 12 and 13). The system's maximum power is measured at four different levels of irradiance at different time intervals in a day.

The comparison of hardware results and simulated results are summarized in Table 3.

The comparison shows that as the solar irradiance is reduced, the maximum power is considerably reduced both in simulation and hardware setup.

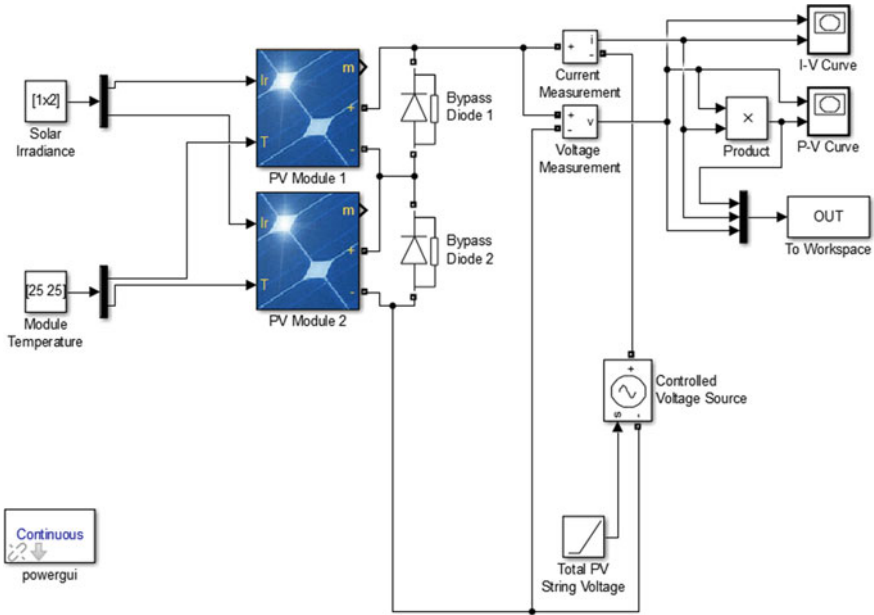


Fig. 3 MATLAB/Simulink model of the PV system

Table 1 PV module specifications

S. No.	Parameter	Value
1	Maximum power	75 W
2	Cells per module	36
3	V_{oc}	21.50 V
4	I_{sc}	4.60 A
5	V_{MPP}	17.60 V
6	I_{MPP}	4.26 A
7	R_s	0.2756 Ω
8	R_{sh}	111.5945 Ω

5 Experimental Validation Using PV Emulator

The concept of irregular solar irradiance is also validated using PV emulator model SPVE001 manufactured by Ecosense Sustainable Solutions Pvt. Ltd. (see Fig. 14). The input parameters of the PV emulator were set as described in Table 4 and the parameters MPP, V_{MPP} , I_{MPP} , FF, and η were noted as shown in Table 5. $I-V$ and $P-V$ curves in four cases of irradiance using PV emulator are shown (see Figs. 15 and 16).

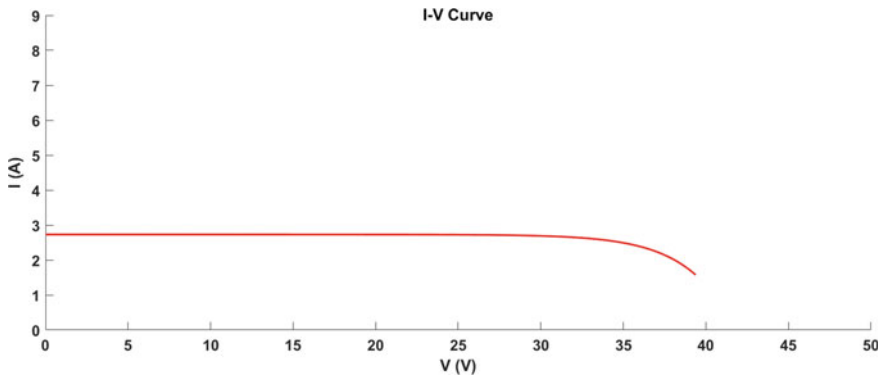


Fig. 4 *I-V* curve in case 1

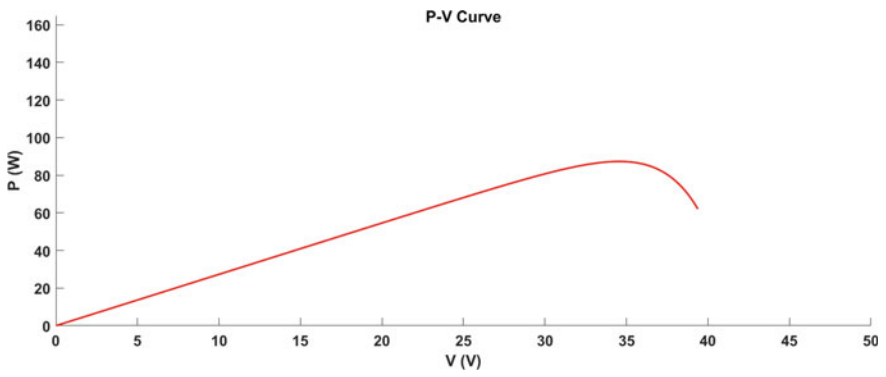


Fig. 5 *P-V* curve in case 1

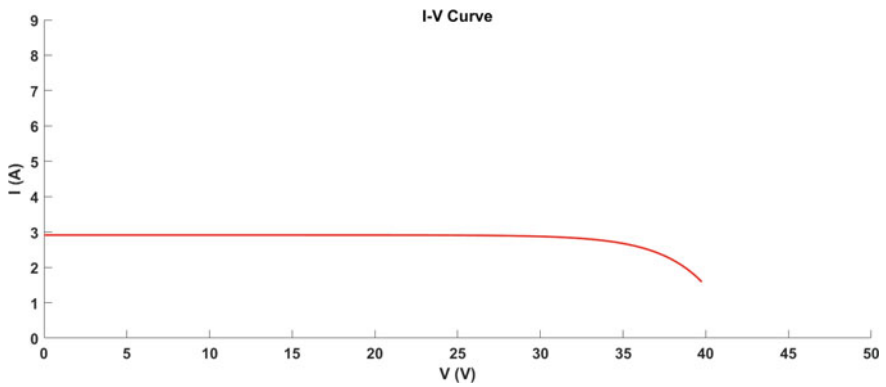


Fig. 6 *I-V* curve in case 2

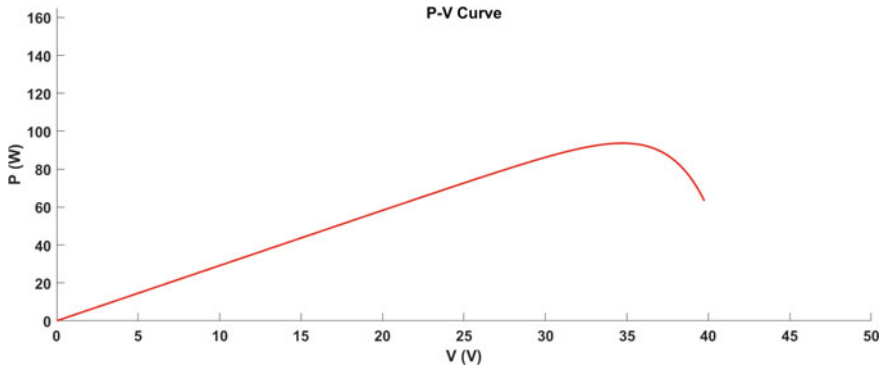


Fig. 7 *P-V* curve in case 2

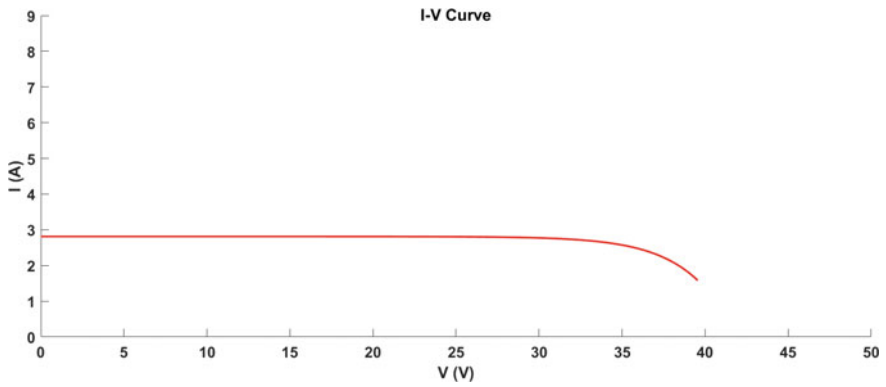


Fig. 8 *I-V* curve in case 3

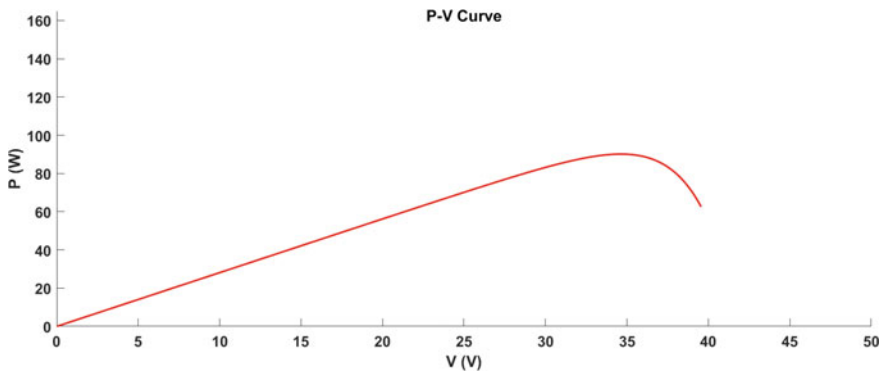


Fig. 9 *P-V* curve in case 3

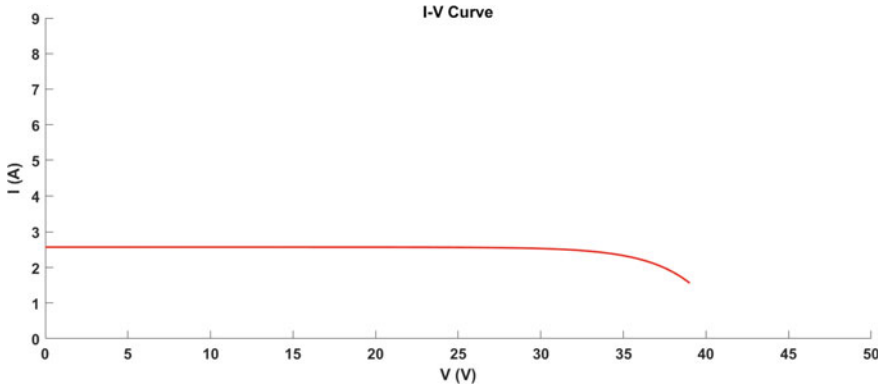


Fig. 10 *I-V* curve in case 4

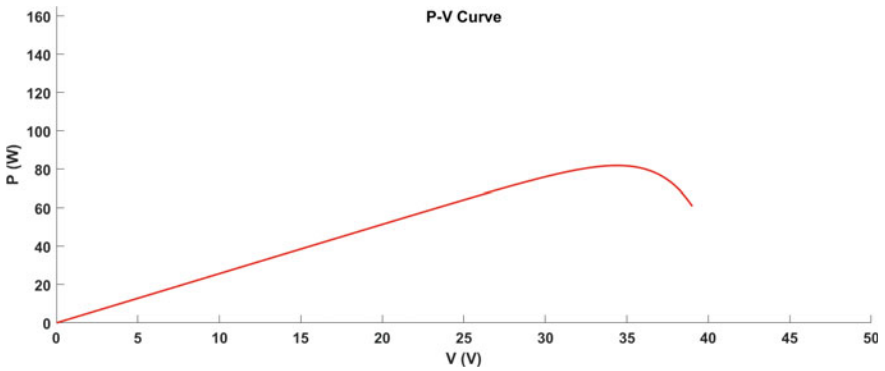


Fig. 11 *P-V* curve in case 4

Table 2 *PV* system performance under four different irradiance levels

Case	Solar irradiance on module 1 (W/m ²)	Solar irradiance on module 2 (W/m ²)	Maximum power (W)
1	641	641	95.05
2	684	684	99.26
3	660	660	98.32
4	604	604	86.28

6 Conclusions

A broad examination of the simulated results highlights the impact of fluctuated solar irradiance on the *PV* system with an increased number of maxima and decreased



Fig. 12 PV system with two PV modules in series

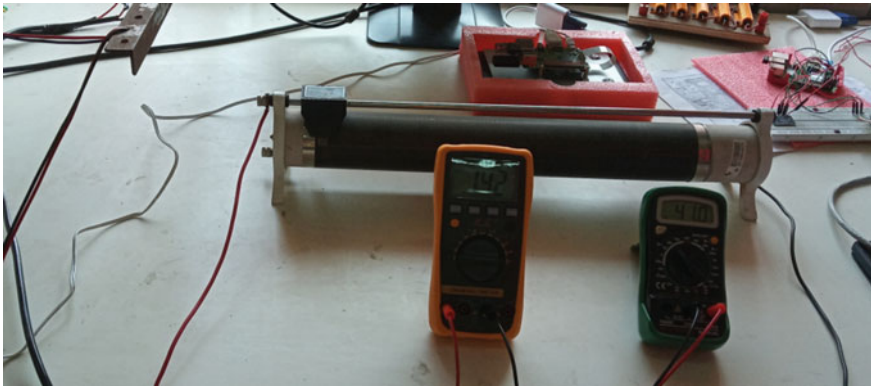


Fig. 13 Test setup for measurement of maximum power

Table 3 PV system performance measured using hardware setup

S. No.	Time	Irradiance (W/m ²)	P_{MAX} (W) using hardware	P_{MAX} (W) using simulation
1	12.15 p.m.	641.5	99.95	95.05
2	12.55 p.m.	683.75	105.55	99.26
3	1.36 p.m.	660	103.14	98.32
4	1.51 p.m.	603.75	91.01	86.28

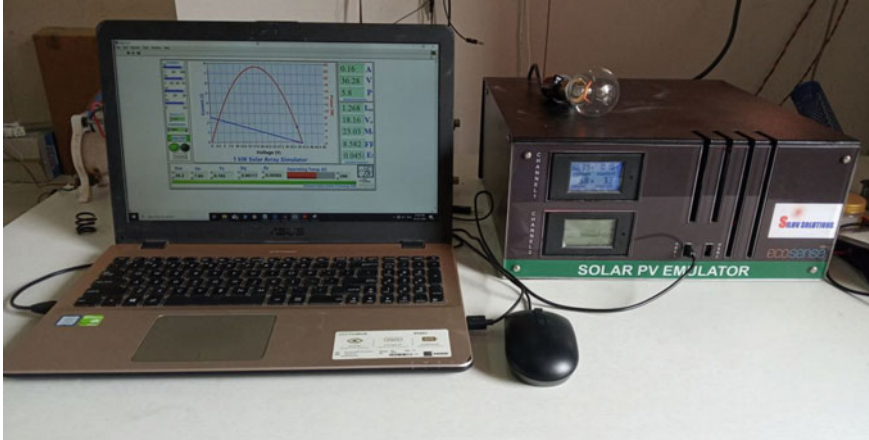


Fig. 14 Experimental setup using PV emulator model SPVE001

Table 4 Input parameters of PV emulator during the experimentation

S. No.	Parameter	Value
1	V_{oc}	36.3 V
2	I_{sc}	7.84 A
3	T_c	0.102
4	a	0.98117
5	R_s	0.039383 Ω

Table 5 Performance parameters of PV array using PV emulator

Case	Irradiance levels (W/m^2)	I_{MPP} (A)	V_{MPP} (V)	MPP (W)	Fill factor	Efficiency (%)	Number of maxima
1	1000, 1000, 1000, 1000	7.32	23.44	171.6	60.32	47.66	01
2	1000, 1000, 1000, 500	7.32	17.58	128.7	46.03	40.85	02
3	1000, 1000, 500, 500	3.755	27.41	102.9	36.82	38.11	02
4	1000, 750, 500, 250	3.755	20.56	77.18	28.11	34.30	04

maximum power. In the first step, experimental validation was done using two series-connected PV modules. A considerable reduction in the maximum power was noticed under changing solar irradiance. The results were also validated using a PV emulator with an analysis of the performance parameters. A considerable reduction in the fill factor, as well as efficiency, was noted.

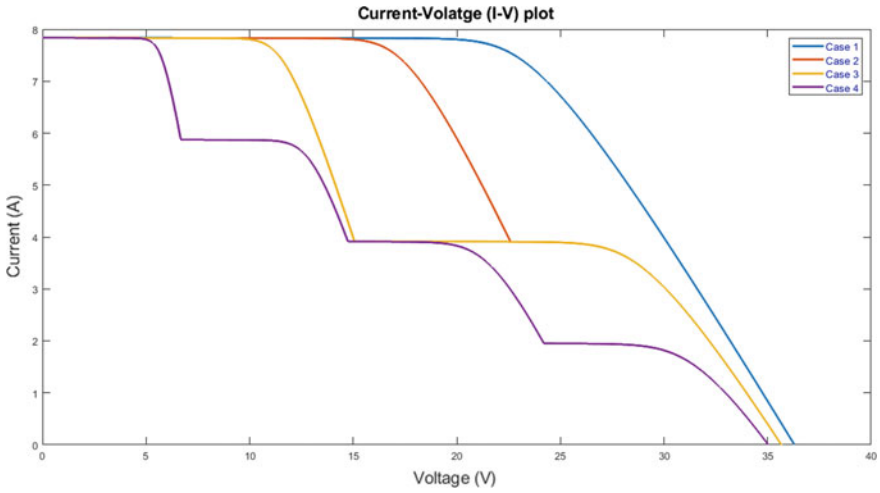


Fig. 15 *I-V* curves in four cases of different irradiance levels using *PV* emulator

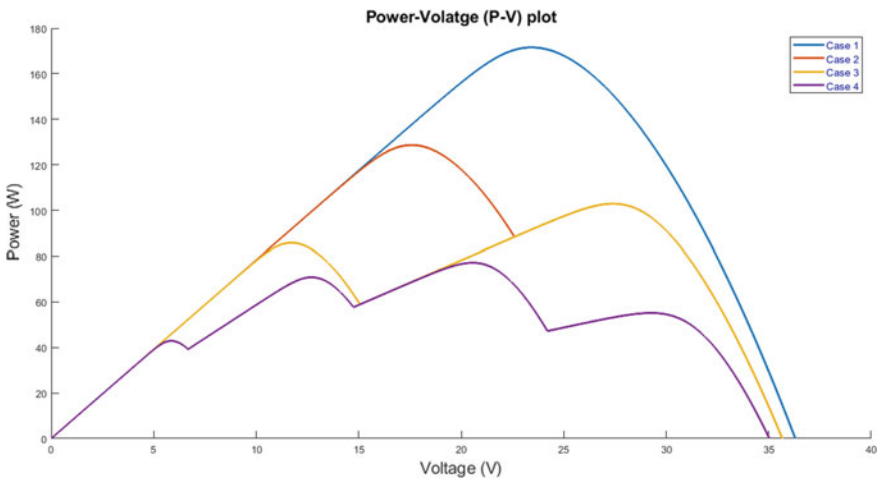


Fig. 16 *P-V* curves in four cases of different irradiance levels using *PV* emulator

The research done in the described work frames an establishment for the judgment about *PV* system execution under irregular solar irradiance. It centers essentially around the requirement for a proficient method to address these damaging components to build a more efficient way of maximum power extraction from the *PV* system.

Acknowledgements *PV* emulator instrument used in the above work is a part of resources available under Technology Business Incubator scheme (Grant No. AR No. 1/Page No. 058/01) funded by

DST, Government of India. The authors would like to thank Symbiosis Institute of Technology, Symbiosis International (Deemed University), Lavale, Pune for this support.

References

1. J.P.N. Torres, S.K. Nashih, C.A.F. Fernandes, J.C. Leite, The effect of shading on photovoltaic solar panels. *Energy Syst.* **9**, 195–208 (2018). <https://doi.org/10.1007/s12667-016-0225-5>
2. A. Belkaid, I. Colak, K. Kayisli, Implementation of a modified P&O-MPPT algorithm adapted for varying solar radiation conditions. *Electr. Eng.* **99**, 839–846 (2017). <https://doi.org/10.1007/s00202-016-0457-3>
3. K. Basaran, Effect of irradiance measurement sensors on the performance ratio of photovoltaic power plant under real operating conditions: an experimental assessment in Turkey. *J. Electr. Eng. Technol.* **14**, 2607–2618 (2019). <https://doi.org/10.1007/s42835-019-00294-8>
4. A. Dolara, G.C. Lazaroiu, E. Ogliari, Efficiency analysis of PV power plants shaded by MV overhead lines. *Int. J. Energy Environ. Eng.* **7**, 115–123 (2016). <https://doi.org/10.1007/s40095-016-0208-2>
5. D.B. Magare, O.S. Sastry, R. Gupta, T.R. Betts, R. Gottschalg, A. Kumar, B. Bora, Y.K. Singh, Effect of seasonal spectral variations on performance of three different photovoltaic technologies in India. *Int. J. Energy Environ. Eng.* **7**, 93–103 (2016). <https://doi.org/10.1007/s40095-015-0190-0>
6. S. Lyden, M.E. Haque, Modelling, parameter estimation and assessment of partial shading conditions of photovoltaic modules. *J. Mod. Power Syst. Clean Energy.* **7**, 55–64 (2019). <https://doi.org/10.1007/s40565-018-0454-9>
7. C. Correa-Betanzo, H. Calleja, C. Aguilar, A.R. Lopez-Nunez, E. Rodriguez, Photovoltaic-based DC microgrid with partial shading and fault tolerance. *J. Mod. Power Syst. Clean Energy* **7**, 340–349 (2019). <https://doi.org/10.1007/s40565-018-0477-2>
8. S. A.L., A.K. Sajina, Maximum Power Point Tracking of PV System under Partial Shading Condition. *Int. J. Eng. Tech.* **4** (2014)
9. F. Salem, M.A. Awadallah, Detection and assessment of partial shading in photovoltaic arrays. *J. Electr. Syst. Inf. Technol.* (2016). <https://doi.org/10.1016/j.jesit.2015.10.003>
10. G. Shankar, V. Mukherjee, MPP detection of a partially shaded PV array by continuous GA and hybrid PSO. *Ain Shams Eng. J.* (2015). <https://doi.org/10.1016/j.asej.2014.10.017>
11. I.H. Mahammed, A.H. Arab, S. Berrah, Y. Bakelli, M. Khennene, S.H. Oudjana, A. Fezzani, L. Zaghba, Outdoor study of partial shading effects on different PV modules technologies. *Energy Procedia* (2017)
12. I.R. Balasubramanian, S.I. Ganesan, N. Chilakapati, Impact of partial shading on the output power of PV systems under partial shading conditions. *IET Power Electron.* (2014). <https://doi.org/10.1049/iet-pel.2013.0143>
13. O. Bingöl, B. Özkaya, Analysis and comparison of different PV array configurations under partial shading conditions. *Sol. Energy* (2018). <https://doi.org/10.1016/j.solener.2017.12.004>
14. R. Ahmad, A.F. Murtaza, H. Ahmed Sher, U. Tabrez Shami, S. Olalekan, An analytical approach to study partial shading effects on PV array supported by literature (2017)
15. R. Boukenoui, H. Salhi, R. Bradai, A. Mellit, A new intelligent MPPT method for stand-alone photovoltaic systems operating under fast transient variations of shading patterns. *Solar Energy* (2016)
16. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
17. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol. 661, ed. by M. Favorskaya, S.

- Mekhilef, R. Pandey, N. Singh (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
18. R.N. Shaw, P. Walde, A. Ghosh, A new model to enhance the power and performances of 4x4 PV arrays with puzzle shade dispersion. *Int. J. Innov. Technol. Explor. Eng.* **8**(12) (2019)
 19. Manjunath, H.N. Suresh, S. Rajanna, Performance enhancement of Hybrid interconnected Solar Photovoltaic array using shade dispersion Magic Square Puzzle Pattern technique under partial shading conditions. *Sol. Energy* (2019). <https://doi.org/10.1016/j.solener.2019.10.068>
 20. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON)* held at Deenbandhu Chhotu Ram University of Science and Technology, SONEPAT, India on Feb 28–March 1, 2020
 21. S. Bana, R.P. Saini, Experimental investigation on power output of different photovoltaic array configurations under uniform and partial shading scenarios. *Energy* (2017). <https://doi.org/10.1016/j.energy.2017.03.139>

Transfer Learning-Based Novel Fault Classification Technique for Grid-Connected PV Inverter



Azra Malik, Ahteshamul Haque, K. V. Satya Bharath,
and Zainul Abdin Jaffery

Abstract The reliability of grid-connected photovoltaic (PV) inverters is of extreme importance and plays a crucial role in maintaining the stability of the grid. In order to prepare the system for different kinds of uncertainties and failures, the identification of the system faults is essential. This paper presents a novel transfer learning-based fault detection and classification technique for grid connected single-phase PV inverters. To achieve this, the voltage and current outputs of the inverter are analyzed by simulating the inverter system for different failure conditions. Further, transfer learning-based classification scheme is utilized for the purpose of fault classification. This novel technique is implemented in MATLAB. Simulation results clearly spell out the efficacy of the proposed technique in terms of accuracy, reliability and robustness.

Keywords Photovoltaic (PV) · Distributed energy resources (DER) · Artificial intelligence (AI)

1 Introduction

Owing to the ever-increasing global energy demand, the reserves of fossil fuels are exhausting at a rapid rate. To relieve the stress on these resources, renewable energy is the preferred solution. Among the available solutions, solar PV is the most abundantly available, clean and environment friendly source of energy [1]. The solar PV systems can be standalone or grid integrated. Grid-connected PV systems utilize inverter as an interface for an efficient power transfer between PV and the grid. Even a minor fault in the inverter may cause imbalance in the system if not handled in

A. Malik (✉) · A. Haque · K. V. Satya Bharath · Z. A. Jaffery
Department of Electrical Engineering, Jamia Millia Islamia, New Delhi, India

A. Haque
e-mail: ahaque@jmi.ac.in

Z. A. Jaffery
e-mail: zjaffery@jmi.ac.in

time. Hence, the study of reliability of inverter is inevitable for ensuring the stability of the overall system. In the event of fault, the system must be capable to identify the fault quickly to avoid further damage [2]. This will save the time required for the repair and maintenance of the affected portion and ensure the continuous operation of the system [3].

Researchers have proposed several fault diagnostic techniques for grid connected PV systems [4–7]. Model-based methods for fault detection are quite popular and are able to achieve considerable positive results [4]. However, these methods are usually time-consuming and complex. Moreover, they are problem specific and cannot be reused for multiple issues. Therefore, the research for alternate solutions is in progress. Among the proposed solutions, data-driven methods have become the major research hotspot these days [5]. Recently, they have gained significant momentum and are used widely for pattern recognition and classification problems [6, 7]. These techniques require signal processing as the first step for useful feature extraction from the fault signals. Different signal processing techniques are available in the literature like Fourier transform, fast Fourier transform and wavelet transform [3]. These techniques are quite complex and time-consuming; therefore, they require significant efforts [8]. Furthermore, the feature extraction process pertaining to a particular fault detection technique remain specific to that problem only and cannot be adapted for a different fault diagnostic technique [9].

In recent times, machine learning [10] and deep learning [9] are a major breakthrough in the area of artificial intelligence (AI). The methods based on machine learning and deep learning are data driven and computationally efficient. They can handle a large amount of data and are considered excellent in the field of pattern recognition and classification. These methods include artificial neural network (ANN) [11], K-nearest neighbor (KNN) method [12], support vector machine (SVM) [13] and naïve Bayes classifier, etc. The accuracy of these methods cannot be further improved because of the shallowness associated with them. Deep neural networks can provide a solution to this [14]. They can automatically extract useful features layer by layer, from the given input. The deep networks take a lot of time for training the data because of the very high numbers of layers associated with them. With the use of transfer learning, it is possible to utilize pre-trained networks with a few modifications in the final layers. It further allows using the already learned features in the pre-trained networks, which helps in optimizing the training time. Therefore, it combines the depth of the deep networks along with reduced training time producing exceptional results. Transfer learning has high prospects of overcoming the shortcomings associated with the traditional machine learning algorithms [15].

This paper proposes a novel fault classification technique for PV inverters. The proposed scheme is able to correctly classify the fault in four different categories and performs exceptionally well with the available data. The technique does not depend upon manual feature extraction from the 1D time series data using time-consuming and complex tools; rather, it utilizes transfer learning as the classification tool, which has automatic feature extraction capability [16]. The deep network utilized in this paper is Residual Network-50 (ResNet-50), which is a 50 layered classic convolutional neural network. This network requires the conversion of 1D time-domain signal

to 2D feature image before feeding it to the network. This conversion takes place through time–frequency analysis of the signal. The methodology adopted for fault identification and classification task and the inference is discussed in the following sections. Section 2 provides the details about the adopted methodology, Sect. 3 is dedicated for simulation results, and the conclusion is provided in Sect. 4.

2 Methodology

2.1 System Structure

The basic structure of two stage single-phase PV inverter system studied is shown in Fig. 1. The system consists of the following parts: data acquisition, 1D to 2D conversion, feature extraction and fault classification using deep neural network. Faults like open circuit faults, short circuit faults and combination of both open and short circuit faults are simulated, and their impact on the inverter output is analyzed. The obtained voltage and current data is used for frequency content visualization and is further converted to 2D spectrograms. These 2D feature images are transferred as input to the ResNet-50 architecture-based network for fault identification and classification.

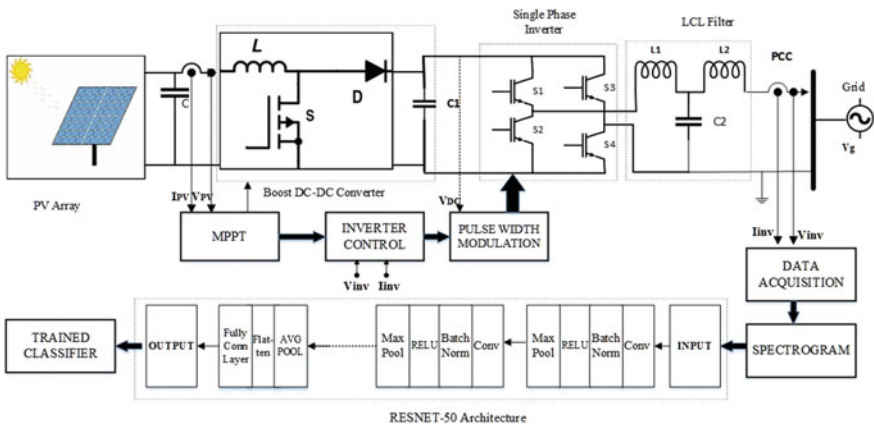


Fig. 1 Grid-connected single-phase inverter system

2.2 *Data Acquisition*

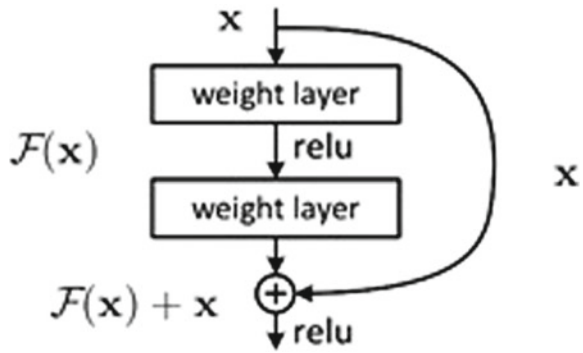
The switching devices are the most likely to encounter undesirable events, and these events have significant impact on the inverter output. Therefore, for smooth and efficient operation of the whole system, it is recommended that the system must be able to identify any kind of fault occurrence. For training the neural network, a large training data is needed. Hence, the data acquisition becomes an important task in order to acquire relevant information regarding various abnormalities occurring in the system.

The inverter output voltage and current output get distorted whenever there is a fault occurrence, and we need to identify and analyze the distortion in order to classify the fault event. Therefore, the inverter switches are subjected to short circuit fault, open circuit fault and combination of open and short circuit fault for the purpose of data acquisition. In each fault case, instantaneous inverter output voltage and current signals in time domain are collected. The acquired data is subjected to 2D conversion with the help of time–frequency details of voltage and current signals. Further, the obtained spectrograms are provided as input to the deep network for minor feature extraction in the initial layers along with the major feature extraction in the deeper layer and fault identification and classification in the final layers.

2.3 *Fault Detection and Classification*

The data acquired at the inverter output is a time series signal. Therefore, the spectrograms are extracted to achieve 2D visual representations of the spectrum at various frequencies. Further, this 2D representation is supplied as input to the transfer learning-based 50-layered ResNet. Transfer learning is a method where an already developed deep network is reused as the starting point for another network [9]. It is a popular approach in deep learning which utilizes pre-trained networks for image classification task. The learned parameters and weights associated with the pre-trained network are adjusted with the new data. The final layers of the network are changed in accordance with the defined task. Transfer learning-based deep networks are preferred these days since they help in optimizing the learning process. They require less time and provide exceptional results compared to other techniques. These networks require large training data, and they do not demand external signal processing since they consist of a layered structure for automatic and powerful feature extraction [16]. The ResNet-50 architecture consists of iterative procedures performed throughout the layers along with the final layer based on the fully connected network architecture [17–20]. These stages are convolutional, pooling and activation layers stacked one over the other. In particular, the convolutional layers perform mathematical operation on the input image data using the filter or kernel and give feature map as the output. The pooling layers identify important features from the feature map matrix and reduce the number of parameters depending

Fig. 2 Skip connection in ResNet



upon the input. The activation function used here is rectified linear activation function (ReLU). The final layer is the fully connected layer, to which we feed the flattened feature matrix, and it uses the softmax function as activation function to provide the classification output [21, 22].

2.4 ResNet Algorithm

Various deep network architectures have been introduced in the previous years like VGG Net, AlexNet and Google Net with the intention to get highly accurate results. All of these have produced groundbreaking results for image classification tasks. It is a well-known fact that with more depth comes tremendously improved accuracy [18]. But there is a limit to this depth also. The backpropagation algorithm calculates the gradients and updates the weights accordingly. As we go into deeper layers, this backpropagation effect causes the initial layers to either learn very slowly or not learn at all. This diminishing gradient problem was a major issue until the introduction of the residual networks. These residual networks tackle the issue of diminishing gradient with the help of skip connections as shown in Fig. 2. The figure describes a residual block having input skipping a few blocks and directly connected to the output. This will aid learning in the initial layers providing outstanding output.

3 Results

The grid connected PV inverter system is developed in MATLAB. It contains a PV array operating under standard operating conditions of temperature 25 °C and irradiance 1000 W/m². The output obtained at the PV panel is enhanced through a boost DC-DC converter. The Perturb and Observe (P&O) maximum power point (MPP) algorithm is employed to extract maximum power from the PV array with the help of boost converter. The output at the DC-DC converter terminals is converted

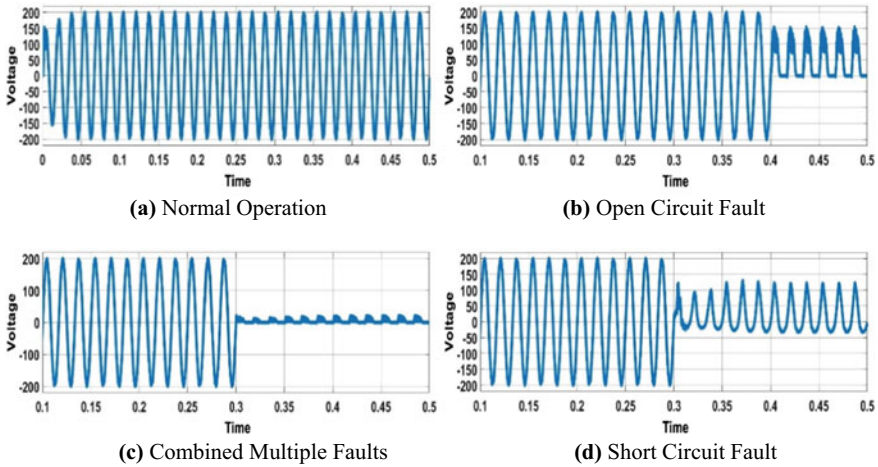


Fig. 3 Inverter output voltage for various faults

to AC through the inverter. A pulse width modulator (PMW) operating at 20 kHz switching frequency supplies gate pulses to the inverter switches. The inverter output corresponds to 230 V RMS with a frequency of 50 Hz. In order to obtain the data for training the classifier, different fault conditions for inverter switches are simulated. These faults include open switch fault, short switch fault and combination of open and short switch faults. Figure 3 depicts the voltage output of inverter for these fault scenarios.

The voltage and current signals at the inverter output are preprocessed and converted to 2D feature images through spectrograms. These feature images are the input into the ResNet-50 convolutional neural network. The network is trained with the ratio of training and validation data as 60:40. The classifier output is collected at the final layer of the network, i.e., softmax classifier. The training and validation results are displayed in Fig. 4. The overall training accuracy is 99%, and the validation accuracy is found to be 97.1%.

4 Conclusion

This paper proposes a transfer learning-based fault classification technique involving deep-layered residual network having powerful inbuilt feature learning characteristic. The ResNet is extremely strong tool for any image classification task compared to other techniques. The proposed methodology is able to identify different kinds of input faults with visibly high accuracy. This will further help in building fault-tolerant systems making them more robust and reliable. These kinds of deep learning-based networks can be applied to other pattern recognition and classification problems since they produce promising results in lesser time and with desired efficiency.

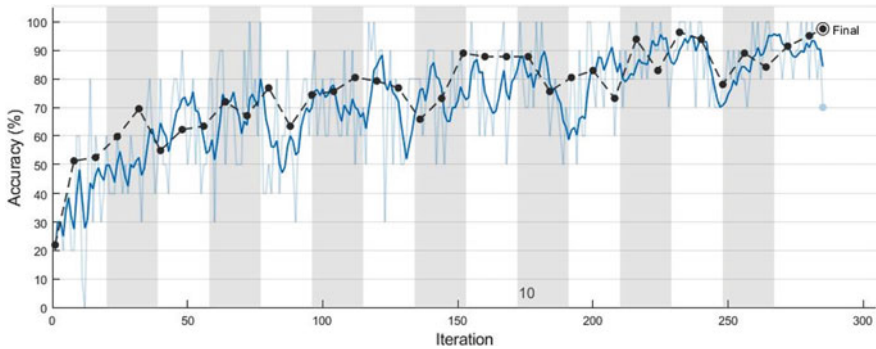


Fig. 4 Classifier training and testing results

References

1. A. Haque, Zaheeruddin, Research on solar photovoltaic (PV) energy conversion system: an overview. *IET Conf. Publ.* **2013**(CP646), 605–611 (2013). <https://doi.org/10.1049/cp.2013.2653>
2. R. Wu, F. Blaabjerg, H. Wang, M. Liserre, F. Iannuzzo, Catastrophic failure and fault-tolerant design of IGBT power electronic converters—an overview. *IECON Proc. (Industrial Electron. Conf.)*, 507–513 (2013). <https://doi.org/10.1109/IECON.2013.6699187>
3. V.S.B. Kurukuru, A. Haque, M.A. Khan, Fault detection in single-phase inverters using wavelet transform-based feature extraction and classification techniques, in *Applications of Computing, Automation and Wireless Systems in Electrical Engineering. Lecture Notes in Electrical Engineering*, vol. 553, ed. by S. Mishra, Y. Sood, A. Tomar A (Springer, Singapore, 2019)
4. J. Poon, P. Jain, I.C. Konstantakopoulos, C. Spanos, S.K. Panda, S.R. Sanders, Model-based fault detection and identification for switching power converters. *IEEE Trans. Power Electron.* **32**(2), 1419–1430 (2017). <https://doi.org/10.1109/TPEL.2016.2541342>
5. V.S. Bharath Kurukuru, F. Blaabjerg, M.A. Khan, A. Haque, A novel fault classification approach for photovoltaic systems. *Energies* **13**(2) (2020). <https://doi.org/10.3390/en13020308>
6. V.S.B. Kurukuru, A. Haque, M.A. Khan, A.K. Tripathy, Fault classification for photovoltaic modules using thermography and machine learning techniques, in *2019 International Conference on Computer and Information Sciences (ICCIS)*, 1–6. <https://doi.org/10.1109/ICCISci.2019.8716442>
7. M.A. Khan, A. Haque, V.S.B. Kurukuru, Machine learning based islanding detection for grid connected photovoltaic system, in *2019 International Conference on Power Electronics, Control and Automation (ICPECA)—Proc 2019-November(1)*. <https://doi.org/10.1109/ICPECA47973.2019.8975614>
8. W. Gong, H. Chen, Z. Zhang, M. Zhang, H. Gao, A Data-driven-based fault diagnosis approach for electrical power DC-DC inverter by using modified convolutional neural network with global average pooling and 2-D feature image. *IEEE Access* **8**, 73677–73697 (2020). <https://doi.org/10.1109/ACCESS.2020.2988323>
9. S. Samal, S.R. Samantaray, M.S. Manikandan, A DNN based intelligent protective relaying scheme for microgrids, in *2019 8th International Conference on Power Systems (ICPS)*, 1–6 (2019). <https://doi.org/10.1109/ICPS48983.2019.9067600>
10. S. Khomfoi, L.M. Tolbert, Fault diagnosis and reconfiguration for multilevel inverter drive using AI-based techniques. *IEEE Trans. Ind. Electron.* **54**(6), 2954–2968 (2007). <https://doi.org/10.1109/TIE.2007.906994>

11. M. Kordestani, M.F. Samadi, M. Saif, K. Khorasani, A new fault diagnosis of multifunctional spoiler system using integrated artificial neural network and discrete wavelet transform methods. *IEEE Sens. J.* **18**(12), 4990–5001 (2018). <https://doi.org/10.1109/JSEN.2018.2829345>
12. F. Harrou, B. Taghezouit, Y. Sun, Improved kN-Based monitoring schemes for detecting faults in PV systems. *IEEE J. Photovoltaics* **9**(3), 811–821 (2019). <https://doi.org/10.1109/JPHOTOV.2019.2896652>
13. C. Bowen, T. Wei, Switch open-circuit faults diagnosis of inverter based on wavelet and support vector machine, in *2019 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, 1178–1184 (2019). <https://doi.org/10.1109/ICEMI46757.2019.9101567>
14. W. Xu, Y. Wan, T.Y. Zuo, X.M. Sha, Transfer learning based data feature transfer for fault diagnosis. *IEEE Access* **8**, 76120–76129 (2020). <https://doi.org/10.1109/ACCESS.2020.2989510>
15. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol. 661, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh, Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-4692-1_58
16. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
17. G. Kapoor, R.N. Shaw, A MMG indices-based scheme for fault detection in series-compensated power transmission system, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, 2020, pp. 742–747 (2020). <https://doi.org/10.1109/ICCCA49541.2020.9250714>
18. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON)* held at Deenbandhu Chhotu Ram University of Science and Technology, SONPAT, India on Feb 28–March 1, 2020
19. J.J.Q. Yu, Y. Hou, A.Y.S. Lam, V.O.K. Li, Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks. *IEEE Trans. Smart Grid* **10**(2), 1694–1703 (2019). <https://doi.org/10.1109/TSG.2017.2776310>
20. M. Haque, M.N. Shaheed, S. Choi, Deep learning based micro-grid fault detection and classification in future smart vehicle, in *2018 IEEE Transportation Electrification Conference and Expo (ITEC)*, pp. 201–206 (2018). <https://doi.org/10.1109/ITEC.2018.8450201>
21. K. Zhang, M. Sun, T.X. Han, X. Yuan, L. Guo, T. Liu, Residual networks of residual networks: multilevel residual networks. *IEEE Trans. Circuits Syst. Video Technol.* **28**(6), 1303–1314 (2018). <https://doi.org/10.1109/TCSVT.2017.2654543>
22. K. Zhang, L. Guo, C. Gao, Optimization method of residual networks of residual networks for image classification, in *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 321–325 (2018). <https://doi.org/10.1109/BigComp.2018.00054>

Improvement of Power System Stability Using FACTS Controller



Harsh Mishra, Faisal Raza, and Shagufta Khan

Abstract This report does the modelling and simulation of thyristor controlled series capacitor (TCSC) and static synchronous series compensator (SSSC) for power network stability enhancement and advancement of power-transfer capability. The TCSC and SSSC are series FACTS controller which is adapted to improve the transmission capability of a transmission line. It is inserted in series of a transmission lines, and it infuses a series voltage in the feeder line in the quadrature of the line current. This report demonstrates the performance test of SSSC and TCSC. All results are carried out in MATLAB software.

Keywords SSSC · TCSC · FACTS · Stability

1 Introduction

In India, the electrical energy demand has increased which is a cause of mismatched generation and demand. It has to be matched for better performance of power network [1–5]. Therefore, FACTS controllers brought in to maintain the accuracy of the system. The FACTS devices can regulate single or more other transmission parameters. The selection of the relevant locations of the FACTS controller is necessary to revamp the voltage stability of the network. In addition, these controllers can be utilized for stepping up the system efficiency and steady state of the network. The FACTS controllers are classified as thyristor-based and VSC-based. Due to the advanced technology of power electronics, the VSC-based FACTS controllers are better in performance [6, 7].

The present-day power system is complicated and it is imperative to accomplish the requirement with improved ability to consume the energy being supplied.

H. Mishra (✉)

Department of Electrical, Electronics and Communication Engineering, Galgotias University, Greater Noida, UP, India

F. Raza · S. Khan

Department of Electronics and Communication Engineering, Galgotias University, Greater Noida, UP, India

Automation is needed to improve the power quality of the power systems. Stability, security, and power advancement are necessary for power quality advancement [8, 9]. There is a need to manage the reactive power in the system to achieve the best fulfilment of the power network. The new feeder lines and generation centres increase the system’s functioning also wider price constraints. Restriction on the expansion of the network has caused in minimizing instability limit, thereby raising the possibility of voltage collapse. The most relevant reactive power source is the FACTS device. In this report, the effectiveness of TCSC and SSSC is analysed and compared.

2 Fundamental Configuration of SSSC

The SSSC is composed of a voltage-sourced converter and a transformer attached in sequence with a feed line. The SSSC infuses volatile voltages in series of a transmission line. The simplified layout of the SSSC is demonstrated in Fig. 1. The comparable circuit of SSSC is displayed in Fig. 2. The equivalent circuit shows the insertion of series voltage and reactance with the transmission line.

This operating mode imitates series reactive rectification in a controlled manner but gives a broader control range as it can run equitably at capacitive or inductive performing domains as well as it can run just as a voltage source. FACTS devices based on thyristor, like TCSC, are reliant on the line current as they control impedance, but in the case of SSSC, it operates as a voltage source, which is independent of the line current [9, 10].

Fig. 1 Static synchronous series compensator

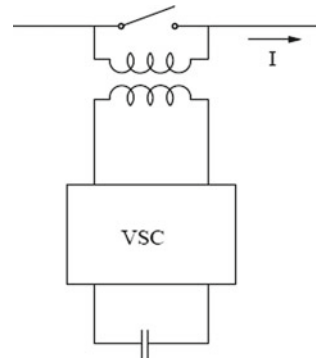
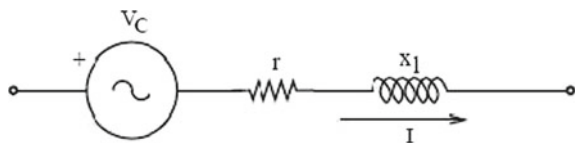


Fig. 2 Equivalent circuit of static synchronous series compensator



The static synchronous series compensator is the most exceedingly impressive FACTS controllers for power flow control. Primary operations in a large area of variable power control are:

- Compensation for long feeder lines.
- Equalization of power flow in lines and staving off of loop flows of active power.
- Receiving end voltage adjustment of a radial line.
- Advancement of transient stability and dynamic stability.

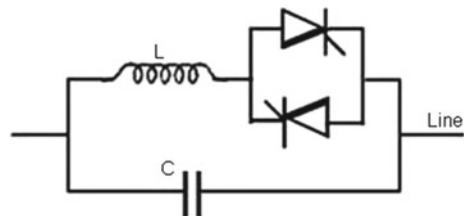
3 Fundamental Configuration of TCSC

TCSC is an excessively necessary series compensator like SSSC. Specifically, in these FACTS device, a thyristor with gate turn-off ability is not needed. Figure 3 displays the simplified diagrams of a TCSC controller.

In thyristor controlled controller, the capacitor is embedded directly into the feeder line, and the thyristor controlled reactor is parallel to the capacitor. As the capacitor is fixed in sequence with the line, thus no requirement to utilize a large potential transformer, and therefore it provides an improved deceleration.

The principle of series compensation directly steps up the frequency voltage across the steady capacitance in a line through the relevant variation of the firing angle. Thyristor controlled controller can be adopted to improve various power network performances such as stability, power oscillations damping, suppress sub-synchronous resonance, and avoid voltage fall down. The efficiency of TCSC controllers is reliant mostly on their suitable placement within the deliberately preferred control signals for accomplishing various functions. Although TCSC functions in highly nonlinear power system surroundings, linear-control methods are practiced largely for the layout of TCSC controllers. During the interruption of a critical line in a connected network, a huge amount of power is likely to flow in parallel transmission paths, which may become firmly overloaded. Contributing fixed-series rectification on the parallel path to develop the power-transfer efficiency emerges to be a practical solution, but it may add total system losses. Therefore, it is beneficial to install a TCSC in key transmission paths, which can accustom its series-compensation level to the instantaneous system needs and grant a lower loss substitute to fixed-series compensation.

Fig. 3 Thyristor controlled series capacitor



4 MATLAB/Simulink Model

4.1 Description of a System

An elementary transmission test system is being used in MATLAB to scrutinize FACTS devices. 11 KV AC is fed. The capacitance of feeder line considered negligible as line is taken short transmission line. Therefore, the inductance is 0.023 mH and the resistance of the line is 5 Ω . The impedance at source is (0.01 + 0001) ohm. The load is kept steady and continual at 25 MV and 50 MVAR. Voltage and current are determined by parameter blocks and active and reactive computed by power measurement block.

4.2 Uncompensated Compensated Model

The MATLAB model of uncompensated system is displayed in Figs. 4, 5, 6, 7, and 8.

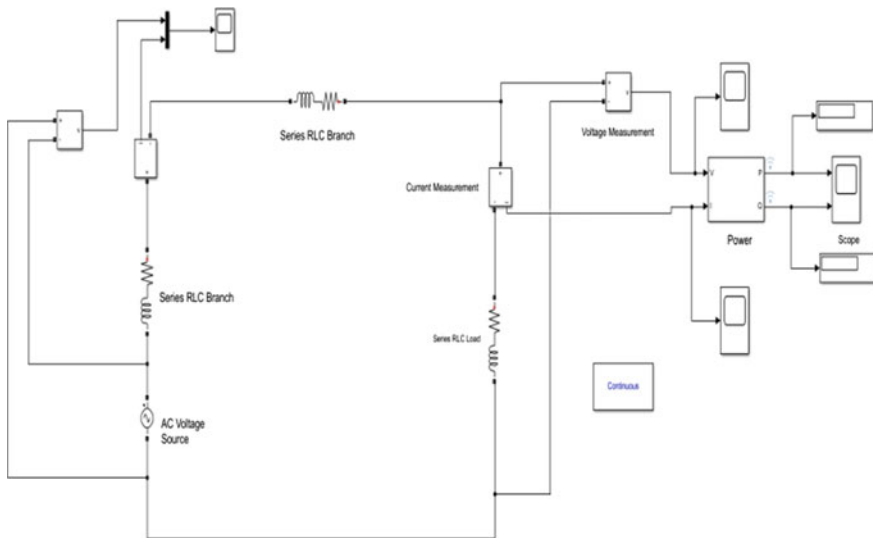


Fig. 4 Uncompensated model

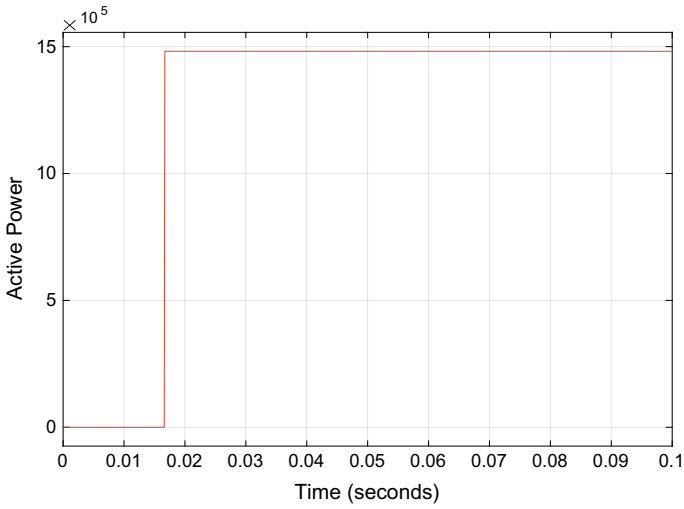


Fig. 5 Active power

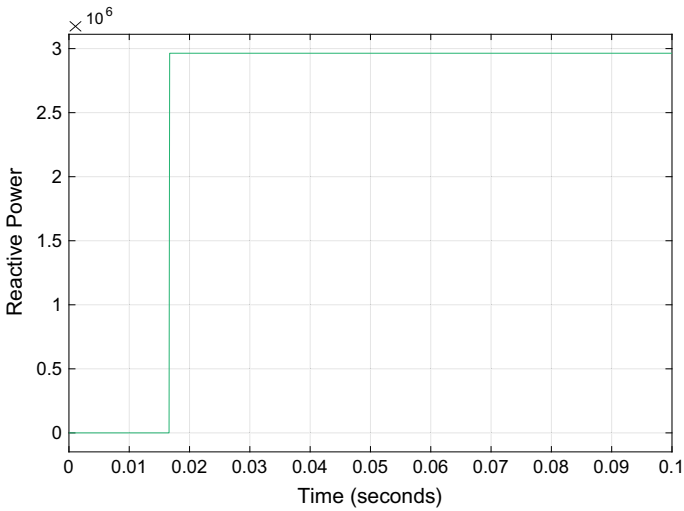


Fig. 6 Reactive power

4.3 TCSC Compensated Model

The model below shows a TCSC attached to the system. In the TCSC model, the inductor is fixed at 100 mH, and outcomes are retrieved for different capacitor values. Outcomes are displayed in Figs. 9, 10, 11, 12, and 13.

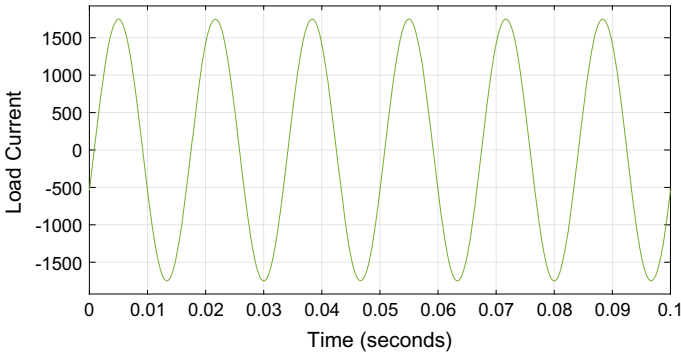


Fig. 7 Load current

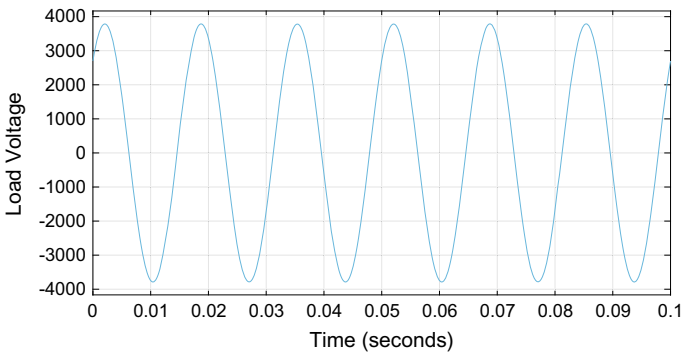


Fig. 8 Load voltage

4.4 SSSC Compensated Model

Graphs for the precise value of capacitance ($350 \mu\text{F}$) are displayed in Figs. 14, 15, 16, 17, and 18. Active and reactive power deviation with the change in capacitance values is formulated in the incoming result section.

5 Simulation Outcomes

5.1 TCSC Compensation

The above graph is sketched when the model is run with a capacitor value of $350\mu\text{F}$. The plots display elevations in load voltage, load current, active power, and reactive

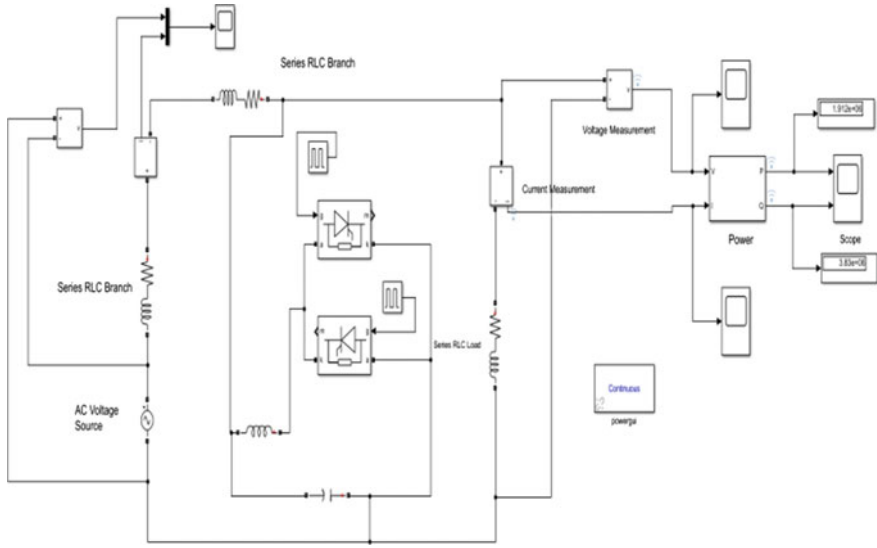


Fig. 9 TCSC compensated model

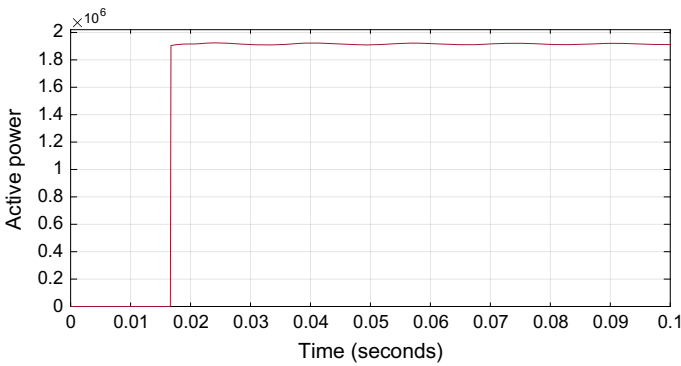


Fig. 10 Active power

power with the integration of TCSC into the system. The results obtained for different capacitor values of the controller are prepared in Table 1.

From Table 1, we can observe that when the capacitance increases, there are stable results of active and reactive power without dissolution. The end voltage rises from 4.2 to 6.3 kV. Voltage develops up to a certain point based on capacitance.

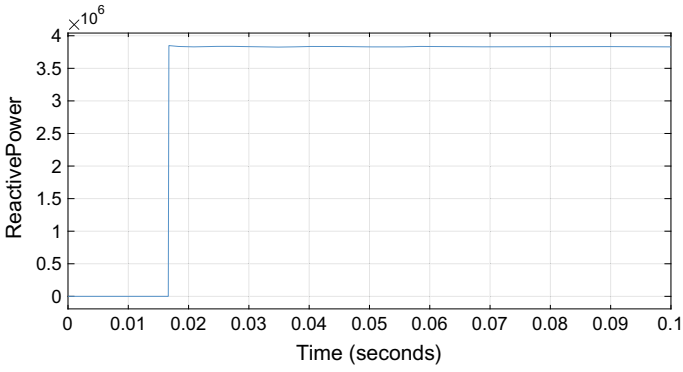


Fig. 11 Reactive power

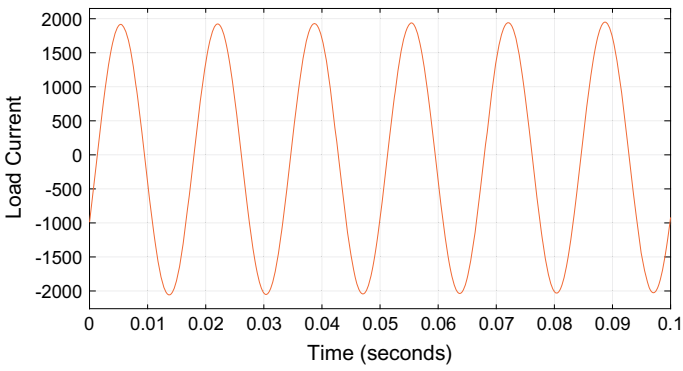


Fig. 12 Load current

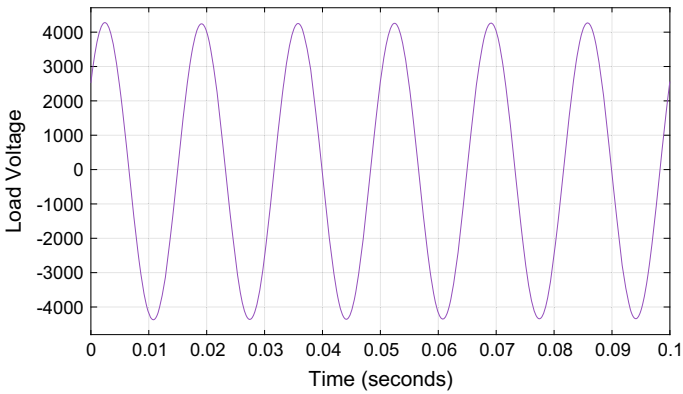


Fig. 13 Load voltage

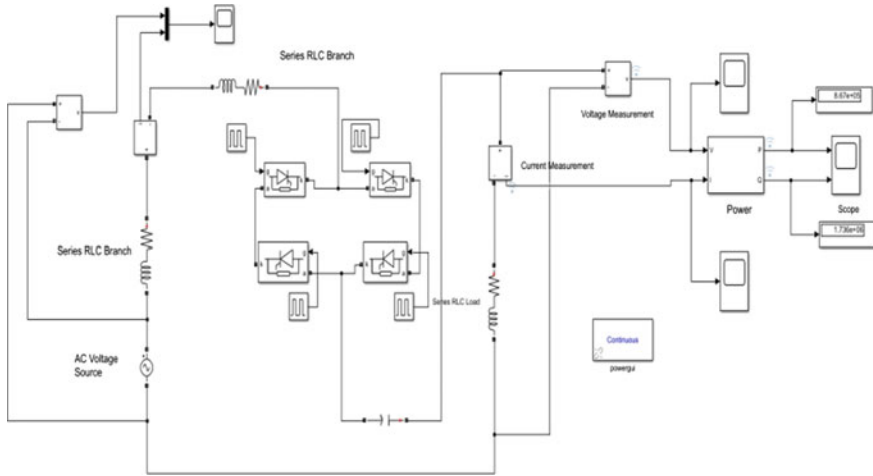


Fig. 14 SSSC compensated model

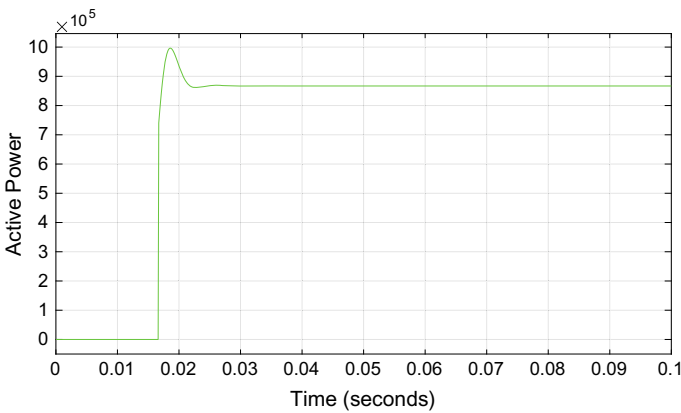


Fig. 15 Active power

5.2 SSSC Compensation

Graphs for the exact value of capacitance ($350 \mu\text{F}$) are displayed in above the section. The change in capacitance along with the active and reactive power variation is formulated in Table 2.

From Table 2, we can check that the active and reactive power increases with capacitance. But, it is also seen that the compensation develops up to a capacity of $600 \mu\text{F}$. If the capacitance is raised beyond this point, both the active and reactive forces are dissolved. So, better compensation is obtained at a capacitance of $600 \mu\text{F}$.

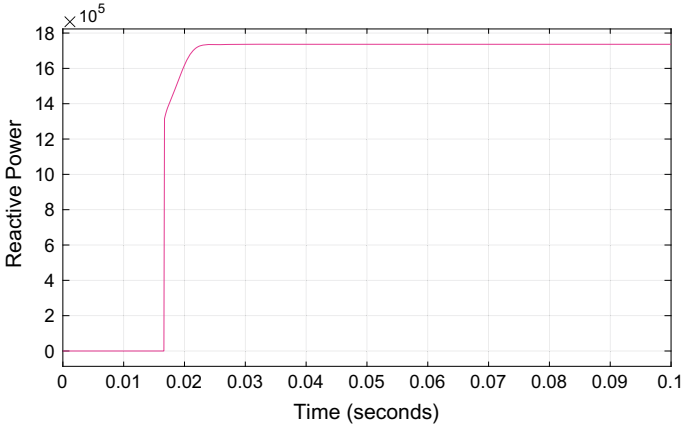


Fig. 16 Reactive power

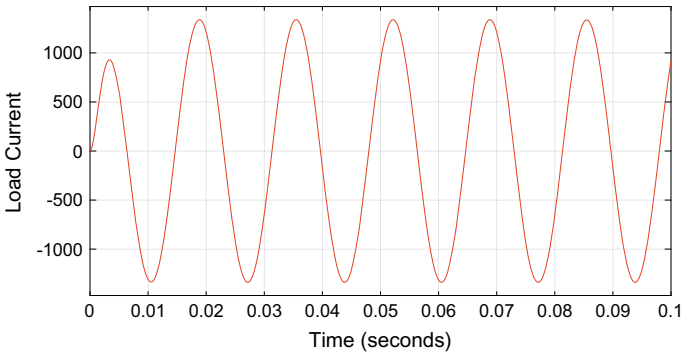


Fig. 17 Load current

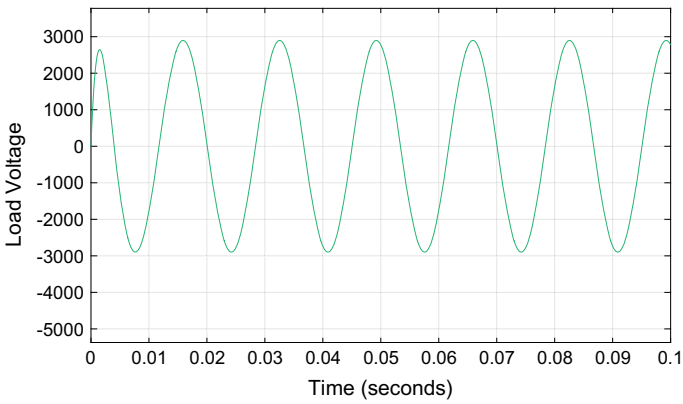


Fig. 18 Load voltage

Table 1 Variation of power at different capacitances

S. No.	Capacitance (μF)	Active power (MW)	Reactive power (MVAR)
1	100	1.52	3.045
2	200	1.668	3.341
3	350	1.912	3.83
4	600	2.35	4.25
5	1000	2.941	5.89
6	1200	3.01	6.032
7	1500	3.12	6.23

Table 2 Variation of power at different capacitances

S. No.	Capacitance (μF)	Active power (MW)	Reactive power (MVAR)
1	100	0.094	0.18
2	200	0.357	0.715
3	350	0.867	1.436
4	600	1.397	2.723
5	1000	1.145	2.123
6	1200	0.833	1.28
7	1500	0.745	1.16

5.3 Comparison Between TCSC and SSSC

Table 3 displays the behaviour of TCSC and SSSC for different capacitor values. However, increasing capacitance indicates the strengthening of a wider volume of equipment. Therefore, after analysing the performance of FACTS devices, it can be seen that better performance is achieved with the inclusion of SSSC in the system for a capacitance of about 600 μF , keeping all other parameters as same.

Table 3 Comparison of power between TCSC and SSSC

FACTS device	Capacitance (600 μF)		Capacitance (1500 μF)	
	Active power (MW)	Reactive power (MVAR)	Active power (MW)	Reactive power (MVAR)
TCSC	2.35	4.25	3.12	6.23
SSSC	1.397	2.7	0.44	1.16

6 Conclusion

MATLAB is practiced for a provisional study to model and simulate TCSC and SSSC attached to a simple feeder line. This report shows a study of FACTS devices and their comparison. Power and voltage are observed to revamp with FACTS devices. Outcomes demonstrate TCSC compensation, reactive power flow revamps gradually with rising capacitor value and is maximal at 1500 μF capacitance. For SSSC compensation, 600 μF capacitance capitulates the most satisfactory outcomes.

But if the capacitance increasing then the price of the apparatus is also rise. Thus, it can be observed that static synchronous series compensator gives the very enticing performance when associated with the network as examine in contrast to TCSC.

Acknowledgements The authors would like to thank the School of Electrical, Electronics and Communication Engineering, Galgotias University, for providing premises for this study.

References

1. S. Akter, A. Saha, Modelling, simulation and comparison of various FACTS devices in power system. *Int. J. Eng. Res. Technol.*, 2–12 (2012)
2. B.R. Andersen and S.L. Nilsson, *Flexible AC Transmission Systems* (Springer International Publishing, 2020)
3. M. Khederzadeh, T.S. Sidhu, Impact of TCSC on the protection of transmission lines. *IEEE Trans. Power Deliv.* (2006)
4. M. Eremia, S. Nilsson, Series capacitive compensation, in *Advanced Solutions in Power Systems: HVDC, FACTS, and Artificial Intelligence: HVDC, FACTS, and Artificial Intelligence*, pp. 339–407 (2016)
5. Y. Belkhier, A. Achour, N. Ullah, R.N. Shaw, Modified passivity-based current controller design of permanent magnet synchronous generator for wind conversion system. *Int. J. Model. Simul.* (2020). <https://doi.org/10.1080/02286203.2020.1858226>
6. G. Kapoor, A. Mishra, R.N. Shaw, S. Ajmera, Complex faults categorization technique for series-compensated transmission lines connected with wind-turbines based on DFVHT, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 703–708 (2020). <https://doi.org/10.1109/GUCON48875.2020.9231152>
7. M. Eremia, S. Nilsson, *Advanced Solutions in Power Systems: HVDC, FACTS, and Artificial Intelligence: HVDC, FACTS, and Artificial Intelligence*, ed. by M. Eremia, C.-C. Liu, A.-A. Edris (2016)
8. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON)* held at Deenbandhu Chhotu Ram University of Science and Technology, SONEPAT, India on Feb 28–March 1, 2020
9. C. Rakpenthai, S. Premrudeepreechacharn, S. Uatrongjit, N.R. Watson, An interior point method for WLAV state estimation of power system with UPFCs. *Int. J. Electr. Power Energy Syst.* (2010)
10. K.R. Padiyar, A.M. Kulkarni, *Dynamics and Control of Electric Transmission and Microgrids* (Wiley-IEEE Press, 2019)

Performance Evaluation of Solar PV Using Multiple Level Voltage Gain Boost Converter with C-L-C Cell



B. M. Kiran Kumar, M. S. Indira, and S. Nagaraja Rao

Abstract The paper presents the performance analysis of a Solar Photovoltaic (SPV) system with a multiple gain boost converter. Mathematical model of a single diode-based SPV is developed and simulated. The corresponding plots of PV and IV characteristics for various temperatures and irradiation are given. The current work also presents the results of MATLAB simulation studies of a SPV system connected to an advanced multiple gain boost converter with Capacitor-Inductor-Capacitance (C-L-C) topology. Generation of different levels of high voltage is possible to meet the requirement of any given application, as this converter is modular in nature. The paper also discusses the performance analysis and comparative study of the SPV system connected to a two-level C-L-C, three-level C-L-C and conventional converters. The simulation results show that the C-L-C converter for SPV is efficient with high gain compared to the conventional DC-DC converter. The peak overshoot in the output voltage of a C-L-C converter is substantially lesser than the conventional converter and decreases with increase in the levels. It is also seen that the output voltage gain is 3 and 4 for a two- and three-level C-L-C converter, respectively.

Keywords Boost converter · Irradiation · DC-DC converter

1 Introduction

Energy generated globally from Renewable Energy Sources (RES) is nearly 28% of the total energy generated from various sources [1]. Energy generated by the solar photovoltaic system is the most popular of all available renewable energy sources as it is clean and green with less maintenance and is inexhaustible [2, 3]. The growth rate of Solar Photo Voltaic (SPV) industry is increasing by about 35–40% every year [4]. As per the 2020 data of International Renewable Energy Agency, total worldwide installed capacity of SPV is about 580 GW. The output of SPV system is non-linear

B. M. Kiran Kumar (✉) · M. S. Indira · S. Nagaraja Rao
Ramaiah University of Applied Sciences, Bengaluru 560058, India
e-mail: kirankumar.ee.et@msruas.ac.in

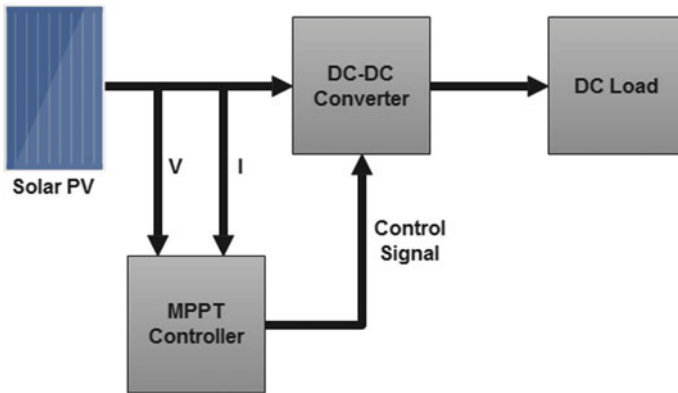


Fig. 1 A typical solar PV system with DC-DC converter

in nature and depends on solar irradiation and temperature. The average efficiency of SPV technology is more than 25% as per the report [5].

The efficiency of the SPV depends on various parameters, such as irradiation, temperature, dust accumulated on the solar panel and partial shading of the cells. Identifying a suitable Maximum Power Point (MPP) of SPV panel is crucial in any solar PV system. Usually, a suitable MPP Tracking (MPPT) extracts maximum power from the SPV. There are various methods in MPPT that are discussed in [6–9]. MPPT is not within the scope of current discussion. A typical 12 V solar PV panel provides a peak voltage of 18 to 21 V [10]. To improve the output voltage, it is essential to have an optimally designed DC-DC converter that can extract maximum power from SPV panel by matching the impedance of SPV to that of the load [11]. Different types of DC-DC converters available in market, such as buck [12, 13], boost [14, 15], buck-boost [16, 17], Cuk [18, 19], SEPIC [20, 21] and zeta [22, 23] are commonly used in SPV systems. A block diagram of a typical SPV system with switched mode power converter is as shown in Fig. 1.

Control of DC-DC converter during the low solar irradianations is a challenging task [24]. The converter duty ratio control helps in matching the SPV impedance with load impedance. A multiple gain boost converter with several C-L-C modules generates high voltages at the output.

2 Mathematical Model of Solar PV

A single current source connected to a diode in parallel represents an ideal PV cell. The practical model has resistances connected in both series and parallel with the model of an ideal cell is as shown in Fig. 2.

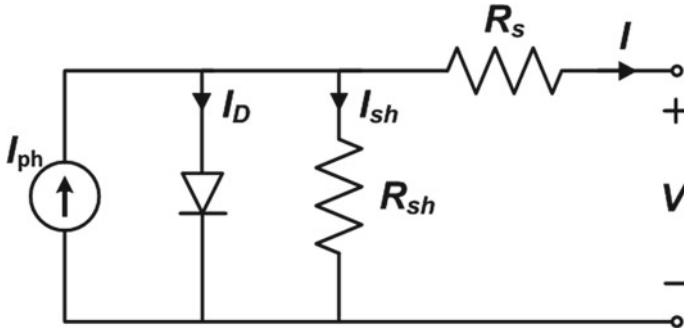


Fig. 2 SPV-Equivalent circuit

2.1 The Electrical Equivalent Circuit-Based Mathematical Model of SPV

The current at the output is given as

$$I = I_{ph} - I_D - I_{sh} \tag{1}$$

$$I_{ph} = [I_{sc} + K_i(T - 298)] \frac{G}{1000} \tag{2}$$

Here, I_{ph} , I_D , I_{sh} and I_{sc} are SPV current, diode current, current through R_{sh} , and the short-circuit current (A), respectively. K_i the short-circuit current of cell 1000 W/m^2 and 25°C , T is the operating temperature in $^\circ\text{C}$ and G the solar irradiation.

Equation (3) gives the diode current I_D in the equivalent circuit

$$I_D = I_o \left[e^{\frac{q(V+IR_s)}{nkTN_s}} - 1 \right] \tag{3}$$

Here, ‘ q ’ the electric charge = 1.6×10^{-19} (C), ‘ n ’ the diode ideality factor = 1.3, ‘ R_s ’ the series resistance (Ω), ‘ k ’ is Boltzmann constant = 1.38×10^{-23} J/K and ‘ N_s ’ the number of series connected cells.

From Eq. (3) the diode saturation current (I_o) given by

$$I_o = I_{rs} \left(\frac{T}{T_n} \right)^3 . e^{\left[\frac{qE_{go} \left(\frac{1}{T_n} - \frac{1}{T} \right)}{nk} \right]} \tag{4}$$

Here, ‘ T_n ’ is nominal temperature, 298 (K) and ‘ E_{go} ’ the bandgap energy of the semiconductor (eV) = 1.1 eV.

And the cell reverse saturation current (I_{rs}) in Eq. (4) is given by

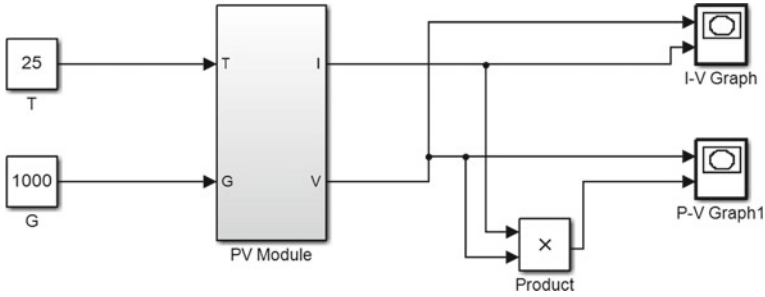


Fig. 3 MATLAB model of SPV

$$I_{rs} = \frac{I_{sc}}{\left[e^{\left(\frac{qV_{oc}}{nkTN_s} \right)} - 1 \right]} \tag{5}$$

Equation (6) gives the value of the shunt current of SPV model

$$I_{sh} = \frac{V + IR_{sh}}{R_{sh}} \tag{6}$$

In Eqs. (5) and (6), V_{oc} is the open-circuit voltage in volts and R_{sh} is Parallel/Shunt resistance in Ω .

For simulation of SPV in MATLAB/Simulink platform, the standard electrical parameters of SPV cell from the manufacturer is necessary. The SPV is modelled based on Apollo ASP4FC200 solar panel ($P_{max} = 200$ W, voltage at $P_{max} = 26.7$ V, current at $P_{max} = 7.5$ A, $V_{oc} = 32.6$ V and $I_{sc} = 8.3$ A). The MATLAB model of SPV is as shown in Fig. 3.

IV characteristics of SPV for various input temperature and irradiation are as shown in Fig. 4a, b, respectively. PV characteristics of SPV for different input temperature and irradiation are as shown in Fig. 5a, b, respectively.

3 Multiple Gain Boost Converter with C-L-C Cell

The output voltage of a solar PV is typically less than its rated value and the output depends largely on temperature and solar irradianations that are intermittent in nature. Hence, direct connection of the SPV output to the load is not possible. Moreover, certain applications require higher voltages than that developed at the converter output. This necessitates use of a boost DC-DC converter to supply a constant high voltage DC output to the load. The boost converter consists of four elements namely, inductor, capacitor, switch and diode to process the power from input to output. A typical boost converter (BC) is as shown in Fig. 6.

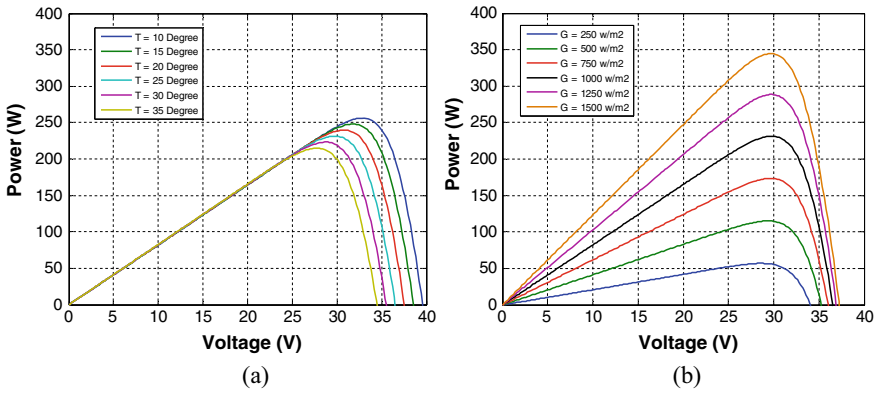


Fig. 4 IV Characteristics of SPV **a** for different temperature, **b** for different irradiation

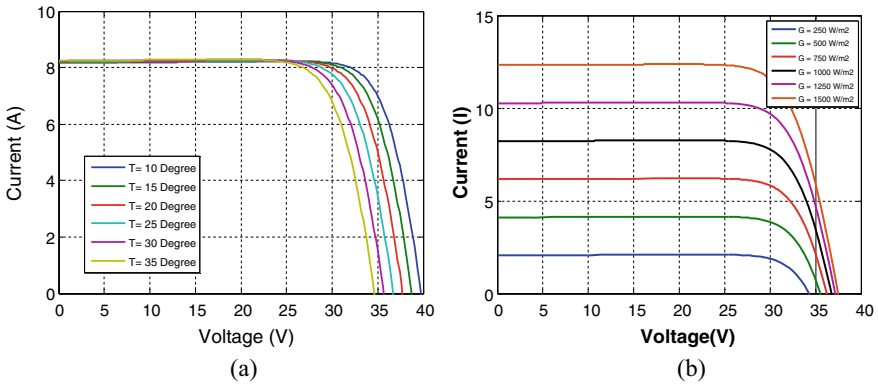


Fig. 5 P-V Characteristics of SPV **a** for different temperature, **b** for different irradiation

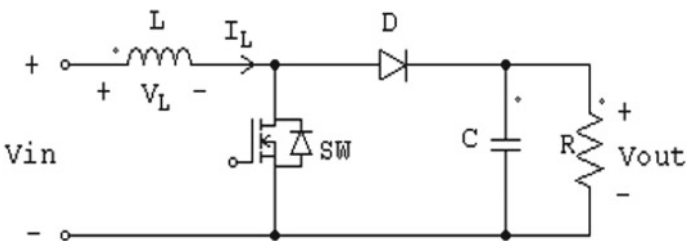


Fig. 6 Conventional BC

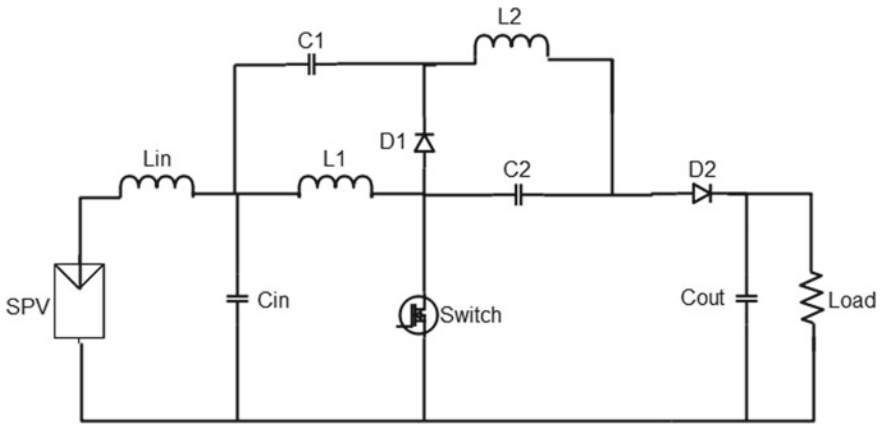


Fig. 7 Two-level voltage lift BC

Regulation of the ON and OFF duration of the switch controls the converter output voltage. The conventional boost converter is bulky and uneconomical, as they require more number of components [25, 26] and large gain being the main limitation. This limitation can be overcome with a C-L-C multiple gain converter [27–30]. The topology is a switched inductor-based boost converter. Each C-L-C module consists of a capacitor-inductor-capacitor configuration to lift the input voltage to multiple levels with a diode between successive modules.

The number of C-L-C modules determines the multiple lifts/levels. The three-level topology currently discussed generates a gain of 4. The circuit diagram of a two-level lift BC is shown in Fig. 7 while three-level voltage lift BC is shown in Fig. 7b.

4 Operating Modes of Multiple Lift Voltage Gain Boost Converter

The C-L-C configuration-based multiple lift voltage gain boost converter consists of number of inductors and capacitors, based on the voltage lift required [26]. A two-level voltage gain module comprises of two capacitors and two switched inductors as shown in Fig. 8a and a three-level voltage gain module consists of three inductors and three capacitors as shown in Fig. 8b. The boost converter has an inductor-capacitor filter at the input to mitigate the oscillations and overshoot present in the output waveform of SPV. The inductor L_1 contributes to the first level of voltage lift and similarly, the second lift in the voltage by inductor L_2 and capacitors C_1 , and C_2 . Similarly, the third lift by inductor L_3 and capacitors C_3 , C_4 . Thus for every level of voltage lift after L_1 (first lift), there is an inductor and two capacitors as shown in Fig. 8b. Hence, known as C-L-C configuration. The diode D_1 enables switching

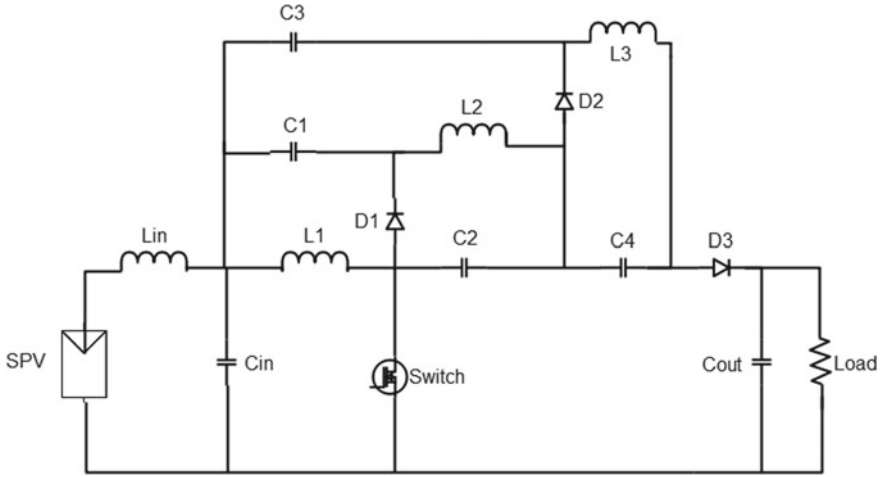


Fig. 8 Three-level voltage lift BC

operation between first and second voltage lift. Similarly, diode D_2 aids the switching operation between second and third level of voltage lift. The diode D_3 turns ON and turns OFF the circuit. A single switch controls the on-off state of the converter. It operates in four different modes based on the on-off status of the switch. Section 4.1 gives the details of the four different modes of converter operation.

4.1 ON State

Mode W: During ON state, the circuit undergoes two different operational modes (Mode W and Mode X). The equivalent circuits for Mode W and Mode X are as shown in Fig. 9a, b, respectively. During Mode W of operation, current at the output of SPV flows through all three levels and while L_1 is charging, L_2 and L_3 discharge, while C_1 and C_3 are charging C_2 and C_4 discharge. Hence, the current in the inductor of the first level increases and currents in inductors of second and third levels decrease. During this mode of operation, the filter capacitor charges and the diodes being reverse biased, the losses associated with the diodes are zero. Mode W operation starts when the current through the inductors in second and third levels decrease to zero.

Mode X: SPV charges the inductors of all three levels during this mode of operation, and hence, the energy is stored in all the inductors and released during the next Mode Y. The currents through the inductors increase in this mode and the direction of the inductor currents in second and third levels will be opposite in direction to that in Mode W. Again, all the diodes in this mode of operation continue to be in a state of reverse bias. The voltage at every level is constant and the relationship between

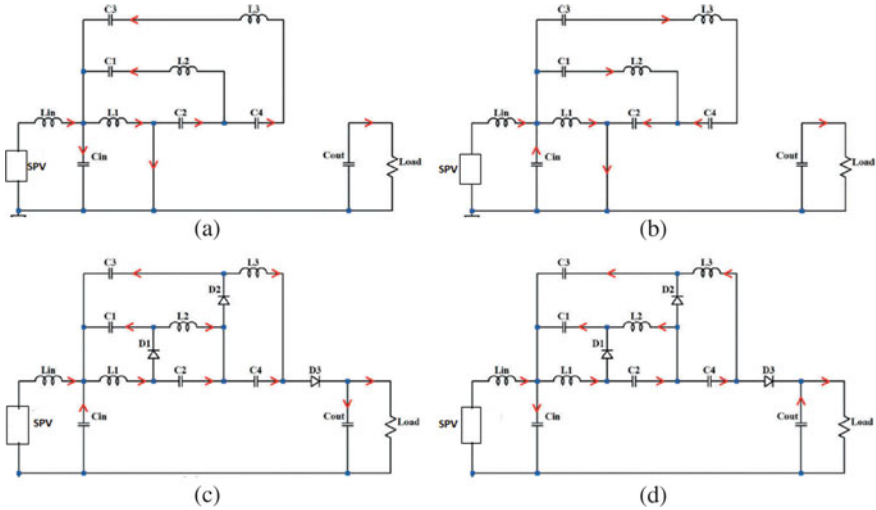


Fig. 9 Equivalent circuits of various operation modes **a** Mode W, **b** Mode X, **c** Mode Y, **d** Mode Z

them is as given in Eq. (7). During the ON state, the converter operates in continuous mode of conduction and the filter capacitor starts discharging.

$$\begin{aligned}
 V_{SPV} &= V_{L1} = V_{L2} + V_{C1} - V_{C2} \\
 &= V_{L3} + V_{C3} - V_{C4} - V_{C2}
 \end{aligned}
 \tag{7}$$

4.2 OFF-State

Mode Y: During ON state of the converter, the circuit undergoes two different modes of operations: Mode Y and Mode Z. The equivalent circuits for these modes are as shown in Fig. 9c, d, respectively. During Mode Y operation, all the diodes are forward biased and current will flow through all diodes. Hence, elements in all levels start discharging and the current flows through the load. In addition, during this mode of operation, inductors satisfy volt-second balance and zero current principle.

Equations (8) to (10) give the voltage across inductors in each level:

$$V_{L1'} = V_{L1} X \frac{D}{1 - D}
 \tag{8}$$

$$V_{L2'} = V_{L2} X \frac{D}{1 - D}
 \tag{9}$$

$$V_{L3'} = V_{L3} X \frac{D}{1-D} \quad (10)$$

$V_{L1'} V_{L2'} V_{L3'}$ = Inductor voltages during off state.

The capacitor voltage in each level will remain constant during all modes of operation.

Equations (11) to (14) give the relationship between inductors and capacitors:

$$V_{C1} = V_{C1'} = V_{L1'} \quad (11)$$

$$V_{C3} = V_{C3'} = V_{L1'} + V_{L2'} \quad (12)$$

$$V_{C2} = V_{C2'} = V_{L2'} \quad (13)$$

$$V_{C4} = V_{C4'} = V_{L3'} \quad (14)$$

The above Eqs. (11) to (14) show that the capacitor voltages C_1 , C_3 and C_2 , C_4 are equal to the voltage across inductors in the preceding lift during OFF state. From the Eqs. (15) to (17), it is clear that the voltages across each inductor are same during the ON state

$$V_{L1} = V_{C1} + V_{L2} - V_{C2} \quad (15)$$

$$= \frac{D}{1-D} X V_{L1} + V_{L2} - \frac{D}{1-D} X V_{L2} \quad (16)$$

$$V_{L1} = V_{L2} = V_{L3} = V_{L4} \quad (17)$$

For a 3-lift/level converter, output voltage is given by sum of SPV voltage and voltage across each inductor as given in Eq. (18).

Equation (19) gives the voltage gain of the converter:

$$V_{out} = V_{in} + V_{L1} + V_{L2} + V_{L3} \quad (18)$$

$$\frac{V_{out}}{V_{in}} = 1 + \frac{3D}{1-D} \quad (19)$$

Mode Z: During this mode of operation the inductor in the first lift continues to discharge and the current flows to the load through the capacitors C_2 and C_4 as shown in Fig. 9d. Inductors in second and third lift along with C_2 and C_4 are charging and the currents being opposite in direction to that in Mode Y. The converter operates in discontinuous mode for small values of duty cycle.

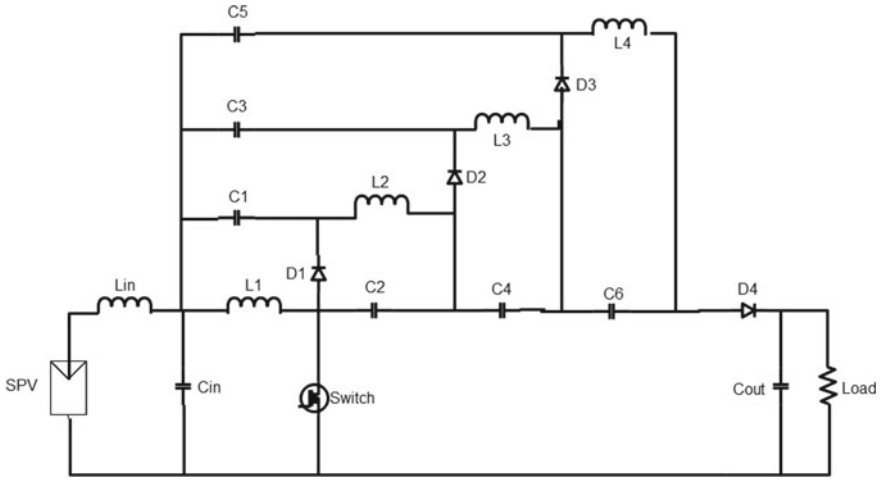


Fig. 10 BC with four- lift C-L-C circuit

Similarly, the operation also extends to ‘*n*’ lifts/levels. Figure 10 shows a four-lift C-L-C circuit for SPV.

The following equations give the voltages for the four-level C-L-C circuit:

$$V_{in} = V_{L1} = V_{L2} + V_{C1} - V_{C2} \tag{20}$$

$$= V_{L3} + V_{C3} - V_{C4} - V_{C2} \tag{21}$$

$$= V_{L4} + V_{C5} - V_{C6} - V_{C4} - V_{C2} \tag{22}$$

The voltage across the inductors in each lift remains the same as discussed in three-lift converter and given by:

$$V_{L1} = V_{L2} = V_{L3} = V_{L4} = V_{L5} \tag{23}$$

Hence,

$$V_{out} = V_{in} + V_{L1'} + V_{L2'} + V_{L3'} + V_{L4'} \tag{24}$$

Equation (25) gives the gain of a four-lift C-L-C converter:

$$\frac{V_{out}}{V_{in}} = 1 + \frac{4D}{1 - D} \tag{25}$$

Similarly, the circuit for multiple lifts with ‘*n*’ lift C-L-C stages is as shown in Fig. 11. It consists of ‘*n*’ number of inductors and capacitors with the output voltage

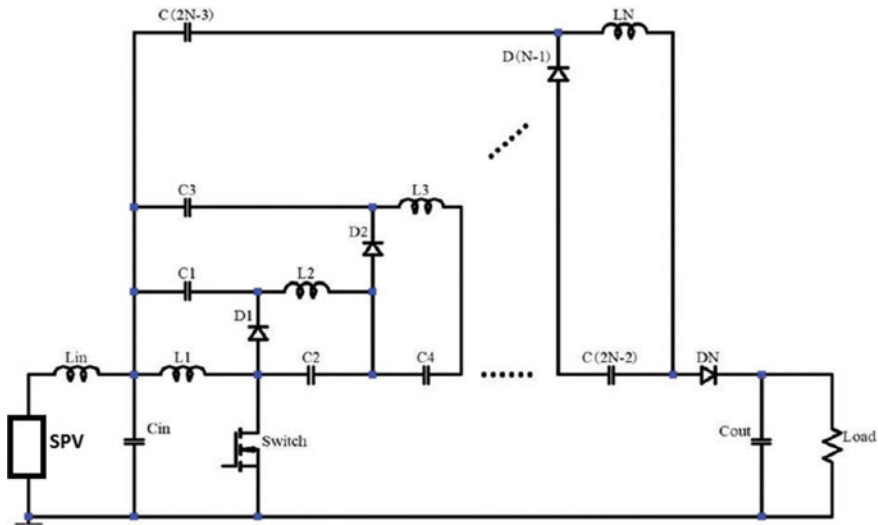


Fig. 11 BC with ‘n’-lift C-L-C circuit

being equal to the sum of SPV voltage across all inductors. The gain of ‘n’ lift C-L-C converter is given by:

$$\frac{V_{out}}{V_{in}} = 1 + \frac{ND}{1 - D} \tag{26}$$

where,

‘n’ = Number of lifts.

5 Results

The MATLAB/Simulink model of SPV with three-lift/level C-L-C boost converter is as shown in Fig. 12 with the output of SPV connected to input of a three-lift BC and converter output fed to a load of 10Ω resistor. The values of L and C of C-L-C circuit are in the range of microhenry and microfarad, respectively. The output performance of C-L-C circuit rates better in comparison with a conventional BC. Output voltage of SPV is 49 V as shown in Fig. 13a. The SPV connected to conventional boost converter with the output measured at 0.5 duty cycle is 100 V as shown in Fig. 13b with a peak overshoot of 82%. A suitable value of filter capacitor at the output side minimizes oscillations in the output of the converter. The same SPV connected to two-lift and three-lift boost converters with 0.5 duty cycle results in an output of 150 V in case of two-lift converter with 64.38% overshoot and an output of 189 V with 43.6% overshoot for three-lift converter as shown in Fig. 13c, d, respectively.

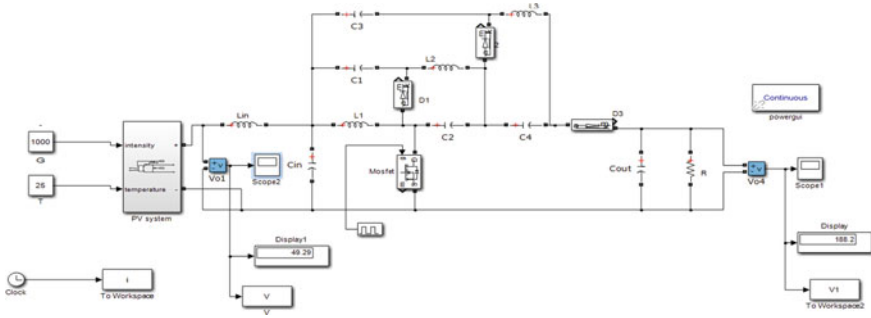


Fig. 12 MATLAB/Simulink model of SPV with three-lift C-L-C converter

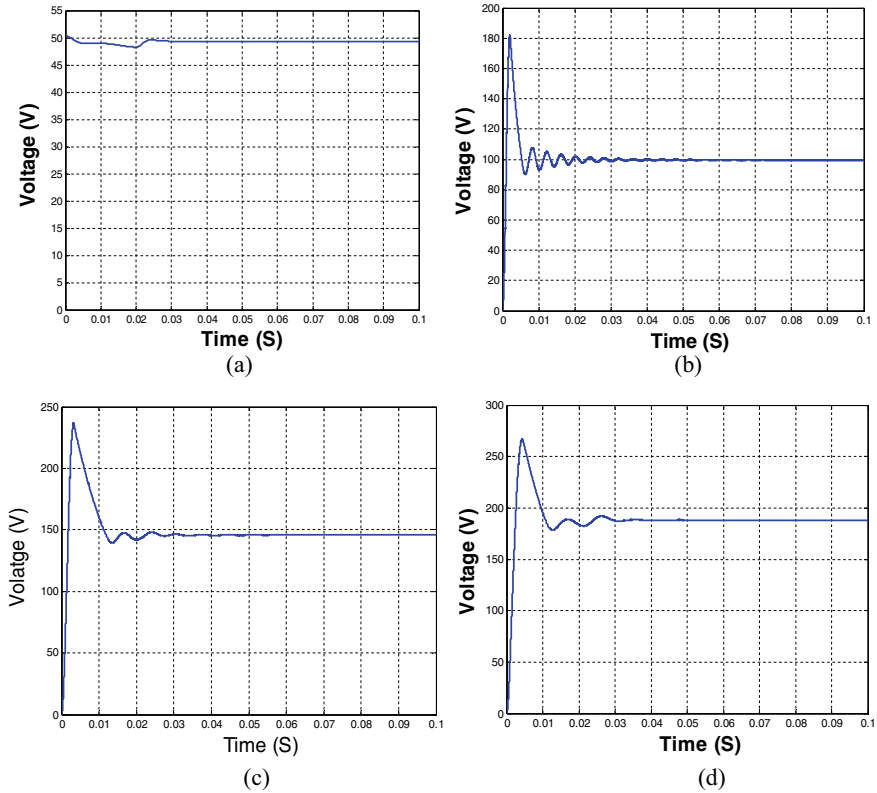


Fig. 13 Output voltage of a SPV, b Conventional BC, c C-L-C two-level shift, d C-L-C three-level shift

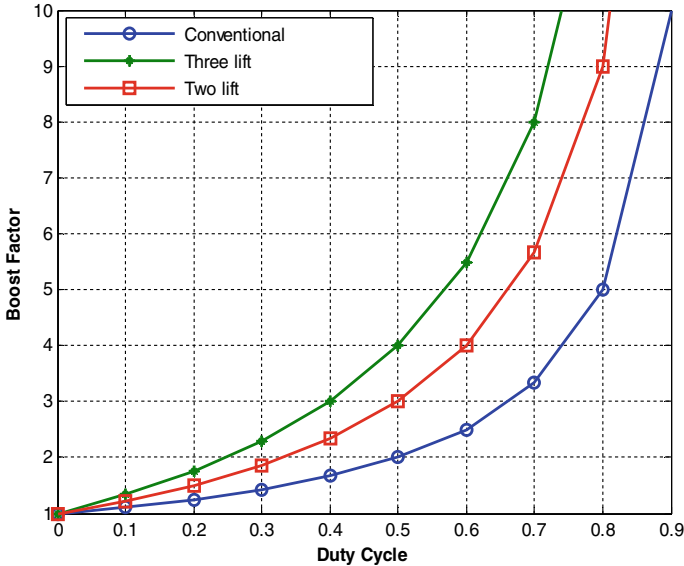


Fig. 14 Boost factor with duty cycle for various lifts

The peak overshoot reduces further with increase in the number of lifts in C-L-C circuit. The output of a three-lift C-L-C converter being four times higher than the input voltage makes it a better choice over a conventional converter. Thus, the gain of C-L-C converter can be adjusted to match the high voltage requirement of any given application. The plot of boost factor versus duty cycle for conventional boost converter and C-L-C converter for two and three lifts are as shown in Fig. 14. For 0.5 duty cycle, the boost factor of conventional converter is 2, for a two- lift C-L-C converter it is 3 and for a three- lift C-L-C converter it is 4.

6 Conclusion

The paper presents the performance analysis of a SPV system connected to multiple gain DC-DC boost converter and mathematical model of a single diode representation of a solar PV with their PV and IV characteristics. The SPV integrated with a C-L-C boost converter produces the required high voltage at the output. This converter being modular in nature is helpful in generating various levels of high voltage, based on the application. The paper also presents the performance analysis of conventional, two-lift and three-lift C-L-C converters along with the comparison of the output waveforms of the converters under discussion for a 0.5 duty cycle. The simulation results show that the C-L-C converter for SPV is efficient with a high gain when compared to a conventional DC-DC converter. The peak overshoot in the output

voltage of a C-L-C converter is substantially lesser than the conventional converter and decreases with increase in the levels. It is also seen that the output voltage gain is 3 and 4 for a two- and three-level C-L-C converter, respectively.

References

1. IEA Homepage, <https://www.iea.org/reports/global-energy-review-2020/renewables>. Last accessed 20 Nov 2020
2. A. Mahesh, K.S. Sandhu, Hybrid wind/photovoltaic energy system developments: Critical review and findings. *Renew. Sustain. Energy Rev.* **52**, 1135–1147 (2015)
3. B.K. Dey, I. Khan, N. Mandal, A. Bhattacharjee, Mathematical modelling and characteristic analysis of Solar PV Cell, in *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pp. 1–5 (2016)
4. H. Patel, M. Gupta, A.K. Bohre, Mathematical modeling and performance analysis of MPPT based solar PV system, in *2016 International Conference on Electrical Power and Energy Systems (ICEPES)*, pp. 157–162 (2016)
5. M. Malinowski, J.I. Leon, H. Abu-Rub, Solar photovoltaic and thermal energy systems: current technology and future trends. *Proc. IEEE* **105**(11), 2132–2146 (2017)
6. K. Yoshikawa, W. Yoshida, T. Irie, H. Kawasaki, K. Konishi, H. Ishibashi, T. Asatani, D. Adachi, M. Kanematsu, H. Uzu, K. Yamamoto, Exceeding conversion efficiency of 26% by heterojunction interdigitated back contact solar cell with thin film Si technology. *Sol. Energy Mater. Sol. Cells* **173**, 37–42 (2017)
7. A.N.A. Ali, M.H. Saied, M.Z. Mostafa, T.M. Abdel-Moneim, A survey of maximum PPT techniques of PV systems, in *2012 IEEE Energytech*, pp. 1–17 (2012)
8. M. Jazayeri, S. Uysal, K. Jazayeri, Evaluation of maximum power point tracking techniques in PV systems using MATLAB/Simulink, in *2014 Sixth Annual IEEE Green Technologies Conference*, pp. 54–60 (2014)
9. Y. Mahmoud, Toward a long-term evaluation of MPPT techniques in PV systems, in *2017 IEEE 6th International Conference on Renewable Energy Research and Applications (ICRERA)*, pp. 1106–1113 (2017)
10. AET Homepage, <https://www.alternative-energy-tutorials.com/solar-power/pv-panel.html>. Last accessed 20 Nov 2020
11. S.N. Singh, Selection of non-isolated DC-DC converters for solar photovoltaic system. *Renew. Sustain. Energy Rev.* **76**, 1230–1247 (2017)
12. N.J. Philips, G.E. Francois, Necessary and sufficient conditions for the stability of buck-type switched-mode power supplies. *IEEE Trans. Ind. Electron. Control. Instrum.* **3**, 229–234 (1981)
13. S. Xu, F. Sun, M. Yang, C. Han, W. Sun, S. Lu, A wide output range voltage-mode buck converter with fast voltage-tracking speed for RF power amplifiers. *Microelectron. J.* **46**(1), 111–120 (2015)
14. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
15. S. Krithiga, N.A. Gounden, Investigations of an improved PV system topology using multilevel boost converter and line commutated inverter with solutions to grid issues. *Simul. Model. Pract. Theory* **42**, 147–159 (2014)
16. R.N. Shaw, P. Walde, A. Ghosh, A new model to enhance the power and performances of 4x4 PV arrays with puzzle shade dispersion. *Int. J. Innov. Technol. Explor. Eng.* **8**(12) (2019)
17. J.W. Yang, H.L. Do, Analysis and design of a high-efficiency zero-voltage-switching step-up DC-DC converter. *Sadhana* **38**(4), 653–665 (2013)
18. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International*

Conference (PIICON) held at Deenbandhu Chhotu Ram University of Science and Technology, Sonapat, India on Feb 28–March 1, 2020

19. C.M. Krishna, Computationally efficient models for simulation of non-ideal DC–DC converters operating in continuous and discontinuous conduction modes. *Sadhana* **40**(7), 2045–2072 (2015)
20. T. Kok Soon, S. Mekhilef, A. Safari, Simple and low cost incremental conductance maximum power point tracking using buck-boost converter. *J. Renew. Sustain. Energy* **5**(2), 023106 (2013)
21. A. Kavitha, G. Uma, M.B. Reesha, Analysis of fast-scale instability in a power factor correction Cuk converter. *IET Power Electron.* **5**(8), 1333–1340 (2012)
22. E. Selwan, G. Park, Z. Gajic, Optimal control of the Cuk converter used in solar cells via a jump parameter technique. *IET Control Theory Appl.* **9**(6), 893–899 (2014)
23. M. Mahdavi, H. Farzanehfard, Bridgeless SEPIC PFC rectifier with reduced components and conduction losses. *IEEE Trans. Industr. Electron.* **58**(9), 4153–4160 (2010)
24. H.L. Do, Soft-switching SEPIC converter with ripple-free input current. *IEEE Trans. Power Electron.* **27**(6), 2879–2887 (2011)
25. S. Narula, B. Singh, G. Bhuvaneswari, Power factor corrected welding power supply using modified zeta converter. *IEEE J. Emerg. Selected Topics Power Electron.* **4**(2), 617–625 (2015)
26. T.F. Wu, S.A. Liang, Y.M. Chen, Design optimization for asymmetrical ZVS-PWM zeta converter. *IEEE Trans. Aerosp. Electron. Syst.* **39**(2), 521–532 (2003)
27. A. Chini, F. Soci, Boost-converter-based solar harvester for low power applications. *Electron. Lett.* **46**(4), 296–298 (2010)
28. N.R. Sulake, A.K. Devarasetty Venkata, S.B. Choppavarapu, FPGA implementation of a three-level boost converter-fed seven-level dc-link cascade H-bridge inverter for photovoltaic applications. *Electronics* **7**(11), 282 (2018)
29. S.N. Rao, D.V. Ashok Kumar, C.S. Babu, Implementation of multilevel boost DC-link cascade based reversing voltage inverter for low THD operation. *J. Electr. Eng. Technol.* **13**(4), 1528–1538 (2018)
30. Y. Li, S. Sathiakumar, J.L. Soon, Multiple lift DC–DC boost converter using CLC cell. *Aust. J. Electr. Electron. Eng.* **16**(1), 46–55 (2019)

Comparative Study for Different Types of MPPT Algorithms Using Direct Control Method



Immad Shams , Karam Khairullah Mohammed , Saad Mekhilef ,
and Kok Soon Tey 

Abstract Maximum power point tracking (MPPT) controllers achieve higher importance due to the high expansion rate of photovoltaic systems. The probability of partial shading conditions (PSCs) is also increased which drastically reduced the maximum power extraction capability from PV systems. In this paper, different types of MPPT algorithms using direct control method have been examined under the same test conditions. The current research trend on MPPT controllers has also been highlighted. This paper gives a brief description to engineers and practitioners working in the area of MPPT controllers.

Keywords MPPT controllers · Partial shading conditions · Direct control method

1 Introduction

Renewable energy resources (RERs) are a major focus of researchers due to the depletion of primary energy sources (such as fossil fuel), global warming, greenhouse effect and alarming pollution rate. The demand of energy is also increasing with the passage of each day due to the development of new technologies such as electric vehicles and increased lifestyle of a common person [1]. Solar energy is said to be the

I. Shams · S. Mekhilef (✉)

Power Electronics and Renewable Energy Research Laboratory, Department of Electrical Engineering, Faculty of Engineering, 50603 Kuala Lumpur, Malaysia
e-mail: saad@um.edu.my

K. K. Mohammed

Control and Mechatronics Engineering Department, School of Electrical Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Malaysia

S. Mekhilef

School of Software and Electrical Engineering, Faculty of Science, Engineering and Technology, Swinburne University of Technology, Melbourne, VIC 3122, Australia

K. S. Tey

Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

253

S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,

Lecture Notes in Electrical Engineering 756,

https://doi.org/10.1007/978-981-16-0749-3_19

prime source of RERs in majority of countries due to its renewability, availability in abundance and environment-friendly nature. These features increased the installation of PV system at standalone and grid level. Solar energy is considered as the next major source to fulfill the increasing demand of energy.

PV module type, maximum power point tracking (MPPT) algorithm, topology utilized for MPPT and converters conversion efficiency are the key factors which affect the PV system performance. MPPT is the most critical point as it is the most economical way to keep the system at its optimum performance. MPPT is a necessary component as well for PV due to the nonlinear nature of power–voltage (P-V) and current–voltage (I-V) curve, as shown in Fig. 1. The MPP varies with the change in the irradiance and temperature condition [2] as shown from the P-V and I-V in Fig. 1, which refer to the uniform shading conditions (USCs).

Conventionally, solar modules are connected in series to provide a high voltage output. 20–24 cells are connected in series composed a module [3]. During partial shading conditions (PSCs) which can be due to dirt, manufacturing error, mismatch in irradiance across each module, etc. The module instead of supplying current will consume which may create hotspot on the surface of the solar panel. To protect the solar modules from hotspot phenomenon, each module is protected with a bypass diode in parallel, which is placed inside the junction box [4]. The activation of bypass diodes will create multiple local maximum power points (LMPP) and one global maximum power point (GMPP) on the P-V and I-V curve, as shown in Fig. 2.

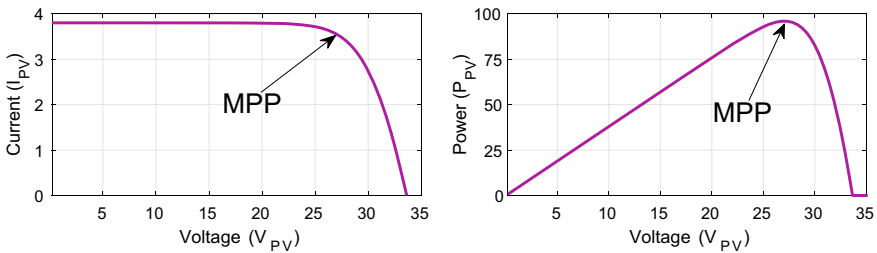


Fig. 1 P-V and I-V curve during USC

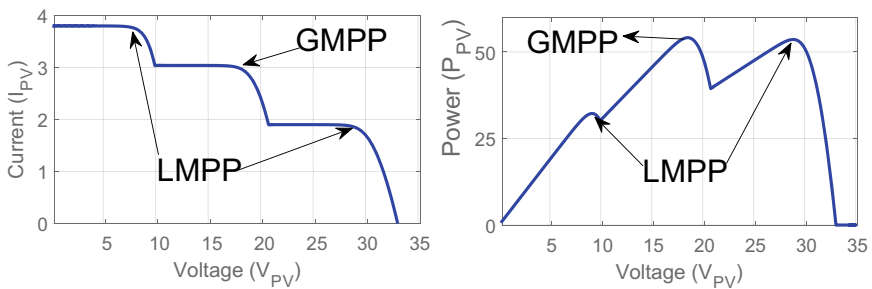


Fig. 2 P-V and I-V curve during PSC

2 Types of MPPT Algorithms

MPPT controllers generally can be divided into conventional, artificial intelligence and metaheuristic-based MPPT algorithms, as shown in Fig. 3.

2.1 Conventional MPPT Algorithms

The conventional MPPT algorithms such as perturb and observe (P&O) and incremental conductance (InC) are very fast in tracking the MPP for USCs. However, when PSCs occurs, the conventional MPPT controllers will be unable to track the GMPP, and they often track LMPP, which will cause power losses. It is due to the way these controllers perform by perturbing the duty cycle and observing the power. Let us see the operating principle of these conventional MPPT controllers.

Perturb and Observe (P&O)

The P&O controller, also known as the hill climbing controller, is one of the most widely used methods for applying MPPT in the PV system [5]. The theory of P&O is to increase or decrease the voltage or duty cycle of the PV array, which has been converted at regular intervals, and to observe the direction of the power change; thereafter, the next control signal has been calculated as shown in Fig. 4. P&O is quick, reliable and fast to enforce. However, the algorithm tracks the MPP by continuously adjusting the terminal voltage of the PV array, which can easily induce output power oscillation. Furthermore, to abrupt environmental changing conditions, the method can result in a power loss of the PV system.

Incremental Conductance (InC)

The InC is another widely used algorithm for MPPT, which works by comparing the instantaneous conductivity I/V to the negative of the slope conductivity dI/dV as shown in Fig. 5, which is equal to the MPP [6]. As environmental factors change, the method will smoothly log changes in the MPP, regardless of the characteristics and parameters of the PV module. However, the InC tracking mechanism is relatively

Fig. 3 Different types of MPPT algorithms

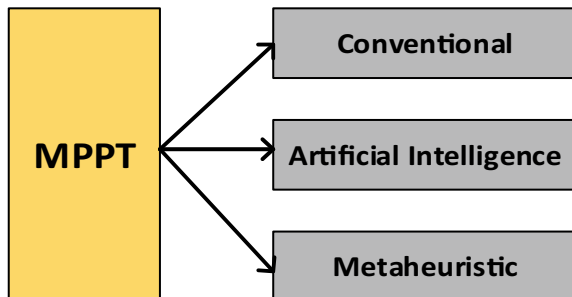
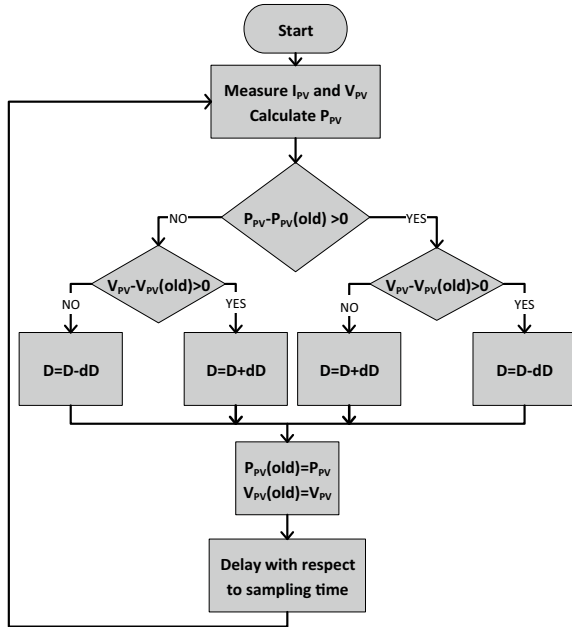


Fig. 4 Flowchart for P&O



complicated and the precision and speed of the measurements will influence the efficiency of the tracking to some level. At the same time, the step size of the voltage increment is also proportional to the error of tracking.

2.2 Artificial Intelligence

Since the conventional MPPT algorithms are unable to track for the PSCs; therefore, artificial intelligence-based MPPT algorithms such as fuzzy logic controller (FLC), artificial neural network (ANN) and adaptive neuro fuzzy interface system (ANFIS) have been proposed to deal with PSCs. However, training the huge amount of data, rule-based table is considered as the biggest drawbacks of such system as it will require high storage. Similarly, they will be unable to track all types of PSCs with high efficiency. Below the basic working operation of these types of MPPT controllers has been elaborated.

Fuzzy Logic Controller (FLC)

FLC is a typical type of artificial intelligence algorithm that is built on a series of fuzzy rules, as shown in Fig. 6. Its execution can be split into fuzzification, assessment of control rules and defuzzification [7]. The core aspect of the FLC is the integration of expert expertise and understanding into the language rules to govern the framework. In addition, FLC can track the MPP easily, and there is no fluctuation in steady-state

Fig. 5 Flowchart for InC

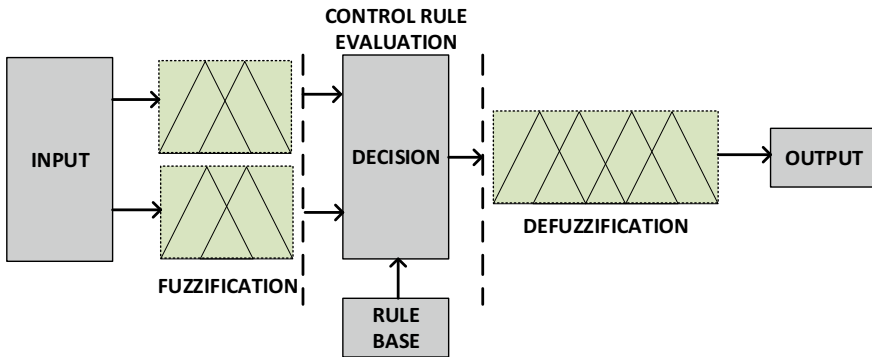
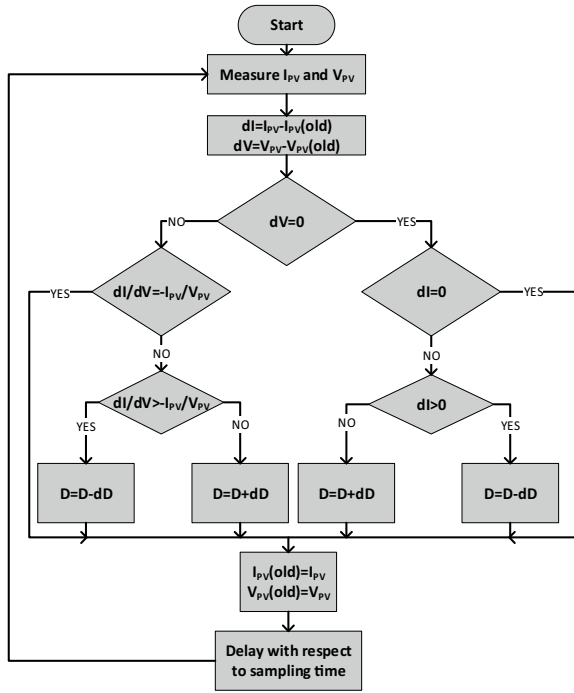
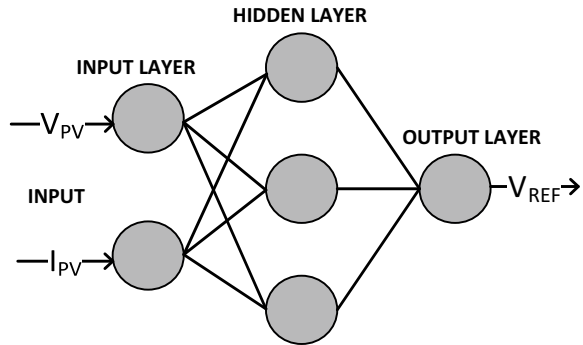


Fig. 6 Block diagram for FLC

conditions after reaching the MPP, i.e., better dynamic and steady-state efficiency. However, the key drawbacks of FLC are the variance in irradiance which can induce drift. The implementation is complex as well. The definition of fuzzy sets, the definition of membership functions and the creation of rule tables require more skill and knowledge from programmers, which directly affects the speed and precision of tracking [8].

Fig. 7 Schematic diagram of ANN



Artificial Neural Network (ANN)

ANN is a modern method of technology for the transmission of information [8]. The most widely used neural network structure has three layers: the input layer, the hidden layer and the output layer, as seen in Fig. 7. When used in the PV system, multiple measurements of input and output data are used to train variables such as network weights and deviations, providing a statistical model of the system's unknown function. The input parameters may be open circuit voltage, short circuit current, illumination, etc., and the output parameters may be duty ratio, voltage, etc. [9]. Training can take time and the number of neurons in the hidden layer has a significant influence on the training time of the network and the precision of the results.

2.3 Metaheuristic-Based MPPT Algorithms

To deal with the PSCs, metaheuristic-based MPPT controllers such as particle swarm optimization (PSO), Jaya, moth flame optimization (MFO), etc., have been proposed which are highly advantageous as they are simple to implement, no requirement of training the data, or study on P-V and I-V characteristic required before its implementation. Metaheuristic-based MPPT algorithms are population-based algorithms in which each particle will move toward the best duty cycle corresponding to the best power. Below the working principle of few of the metaheuristic-based MPPT algorithms has been highlighted.

Particle Swarm Optimization

Particle swarm optimization (PSO) is a population-based metaheuristic algorithm, proposed by Kennedy and Eberhart in 1995 [10], brought a revolution in the field of bio-inspired algorithm. Fish schooling and bird flocking inspired its mathematical modeling. Its metaheuristic approach and capability to solve optimization problems with many local maxima makes its utilization popular. Several cooperative agents are utilized where they share the information obtained in a defined search space. The

number of cooperative agents is assigned in initialization stage. Each agent follows two simple rules, i.e., to follow the local best means agent’s self-best position and follow the global best agent means the best position among all agents as shown in Eq. (1). Ultimately, each agent will evolve around the optimal solution. The PSO method can be defined using standard set of equations as shown in Eqs. (1) and (2).

$$v_i(k + 1) = \underbrace{w \cdot v_i(k)}_{\text{inertial motion}} + \underbrace{C_1 \cdot r_1 \cdot (P_{\text{best},i} - X_i(k))}_{\text{Local influence}} + \underbrace{C_2 \cdot r_2 \cdot (G_{\text{best}} - X_i(k))}_{\text{Global influence}} \quad (1)$$

$$X_i(k + 1) = \underbrace{X_i(k) + v_i(k + 1)}_{\text{Update of agent's position}} \quad (2)$$

The coefficients of C_1 , C_2 and w are of higher importance. The higher coefficient values may result in fast convergence but at the same time may result in wrongly optimized peak point. The smaller values of coefficient will result in slow convergence speed toward optimum point. Hence, proper values selection for these constants is of higher importance. The flowchart for PSO is shown in Fig. 8a.

Jaya

Another population-based powerful global optimization algorithm known as Jaya is presented in [11]. The biggest advantage of Jaya in comparison with other meta-heuristic algorithms is that it has no tuning parameter, which makes the complexity

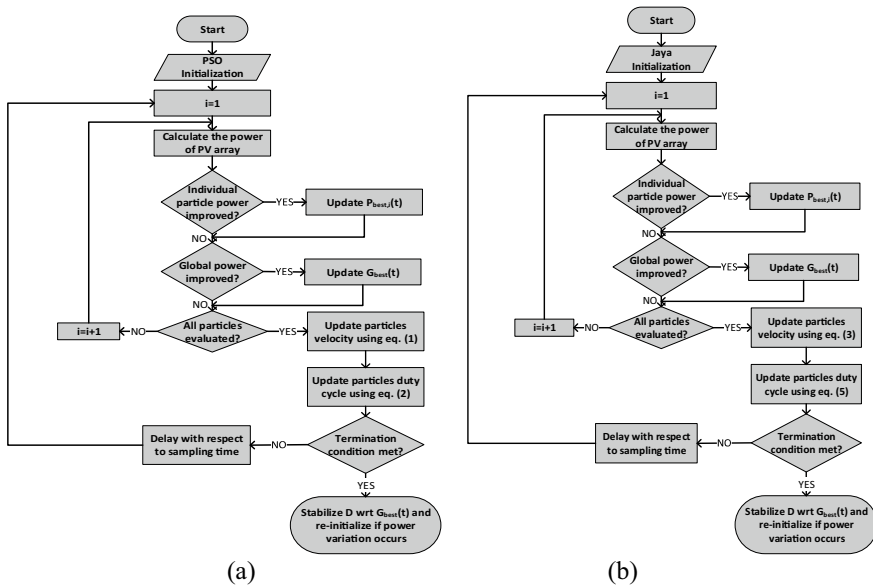


Fig. 8 a Flowchart for PSO and b flowchart for Jaya

much simple. It is based on moving toward the best solution avoiding worst solution as shown in Eq. (3). Then, the updated position is compared with the previous position as in Eq. (4). If it gives better solution, then it replaced with the previous position and vice versa as in Eq. (5). Only population size and termination criteria are required. The flowchart for this method is shown in Fig. 8b.

$$\text{Update: } X_i(k + 1) = X_i + r_1 \cdot (X_{\text{best}} - X_i) - r_2 \cdot (X_{\text{worst}} - X_i) \tag{3}$$

$$\text{if } X_i(k + 1) > X_i \text{ then update the position} \tag{4}$$

$$\text{Else } X_i(k + 1) = X_i \tag{5}$$

3 Comparison

In this section, the SEPIC converter has been utilized with component parameters as shown in Fig. 9. Three PV modules connected in series with each module PV parameters; P_{MPP} as 58 W, V_{MPP} as 12.2 V, I_{MPP} as 4.78 A, V_{OC} as 14.67 V and I_{SC} as 5.17 A. Three different irradiances have been tested including one with USC 1000 W/m² and two with PSCs as 1000, 500, 300 W/m² and 1000, 900, 800 W/m² with constant temperature of 25 °C. The MPP and GMPP for these patterns are shown in Fig. 10a. The comparison has been elaborated between one controller from each type, such as P&O, FLC and PSO, to determine the difference of tracking state between them. For P&O, the dD is selected as 0.01, whereas for PSO, C_1 and C_2 are selected as 1.2 and 1.5, whereas w is selected 0.4. The sampling time chosen for P&O is 1 ms, whereas the sampling time chosen for PSO is 50 ms.

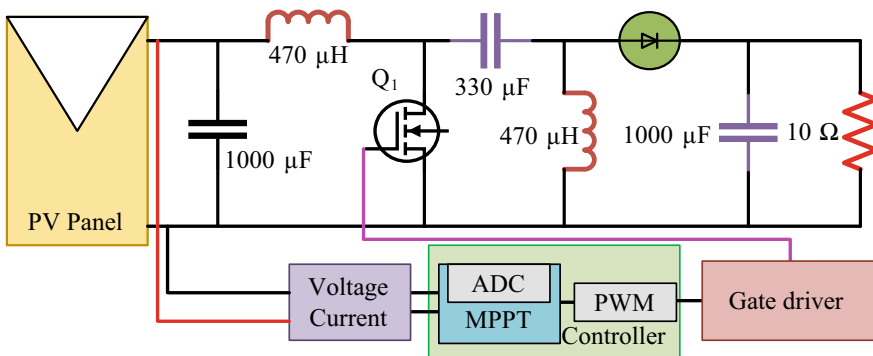


Fig. 9 Block diagram using SEPIC converter for MPPT implementation

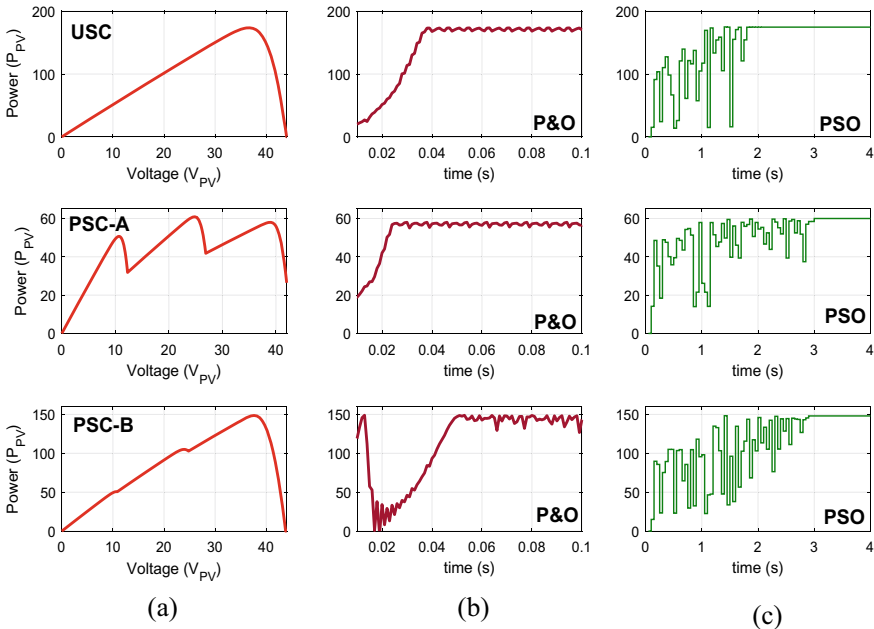


Fig. 10 a Different shading patterns, b response of P&O and c response of PSO

The results in Fig. 10 shows that conventional MPPT algorithms show low oscillation during the tracking state for USC and very fast tracking speed towards MPP. The main reason for rapid P&O tracking is a small sampling time of 1 ms. The small sampling time for P&O is possible due to the small disruption of the duty cycle. However, steady-state oscillation was observed when the MPP was tracked for USC.

For PSO, it can be observed for USC that it can correctly tracks the MPP point, however, the power oscillations are high, whereas the tracking time is also high. It is due to the reason of high sampling time in comparison with P&O which is taken as 50 ms. The sampling time for PSO is always taken higher due to abrupt changes of duty cycle from one point to another in a search space. However, as it can be observed that the steady-state oscillation has been avoided.

It can be observed for P&O that it cannot track the GMPP for PSC-A and get steady-state oscillation at the LMPP. However, the PSO successfully tracks the GMPP for PSC-A. The P&O tracks the GMPP for PSC-B because its initial value of the duty cycle was in that GMPP region, therefore, it managed to track the GMPP.

The FLC and ANN track the MPP for USCs with faster convergence speed then P&O but making a rule-based table for PSCs and training the data for PSCs will be a challenging task and it will not be able to accurately determine the GMPP region for all the PSCs as well. Therefore, it has not been simulated in this paper.

4 Current Research Trend

Currently, a variety of experiments have been undertaken on hybrid-based MPPT algorithms to resolve the consequences of PSCs affecting convergence speed and MPPT performance with less storage and computational pressure.

The conventional P&O algorithm has been modified in [12] to achieve zero oscillation during steady state. InC was modified [13] to work for PSCs with three multiple peaks. An improved InC approach has been suggested in [14], in which an allowable error is made within a certain precision range that satisfies the MPPT. This changed procedure will easily detect the change of irradiance direction and make the correct decision. Similarly, zero steady-state oscillation is reached, which eliminates power losses.

Some researchers have sought to merge the standard MPPT algorithm with FLC to increase the convergence speed and reduce the steady-state oscillation and implementation complexity. The beta variable was implemented in [15], which simplifies the membership function and reduces the reliance on the simple rule architecture. In addition, FLC was also suggested for an MPPT scheme in which four optimization algorithms are provided for the optimization of fuzzy membership functions and the generation of a proper duty cycle [16]. ANN also hybridized with other algorithms to produce better performance. In order to minimize the training time, the PSO is used to maximize the number of RBFNN parameters, change the weight parameters and introduce the phase velocity to achieve the result of the applying regulation [17].

Metaheuristic-based MPPT algorithms are unable to distinguish between solar intensity or load variation, and thus re-initialize from the beginning to detect MPP even though load variation occurs. In [18], a very fast tracking system has been suggested for load variations that will require a maximum of two sampling times to track back the GMPP, while the load varies. However, the controller is system-dependent and would require a modification in the converter. To address the problem of converter dependence in [19], a new approach to load variation has been proposed which avoids both re-initialization and non-converter dependence. Similarly, metaheuristic algorithms are unable to distinguish between PSC and USC, while in [20], a search space skipping approach has been suggested for metaheuristic algorithms to increase tracking time for USCs. A P&O has been hybridized with metaheuristic algorithm in [21], in which P&O will scan the left most peak first to determine whether the PSC is present or not.

5 Conclusion

In this paper, a general comparison has been elaborated and compared between different types of MPPT algorithms. MPPT algorithms have been classified into conventional, artificial intelligence and metaheuristic-based MPPT algorithms. One

controller has been chosen from each type to illustrate their operation through simulation. Lastly, the current research trend on MPPT controller has also been highlighted to determine the direction of research for future.

Acknowledgements The authors would like to thank the Ministry of Higher Education, Malaysia for providing the financial support under the Long Term Research Grant Scheme (LRGS): LRGS/1/2019/UKM/01/6/3.

References

1. R. Ahmad, A.F. Murtaza, H.A. Sher, Power tracking techniques for efficient operation of photovoltaic array in solar applications—a review. *Renew. Sustain. Energy Rev.* **101**, 82–102, 2019/03/01/ 2019
2. A.A. Ramaden, I.A. Smadi, Partial shading detection and global MPPT algorithm for PV System, in *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*, 2019, pp. 135–140
3. K.A.K. Niazi, Y. Yang, D. Sera, Review of mismatch mitigation techniques for PV modules. *IET Renew. Power Gener.* **13**, 2035–2050 (2019)
4. K.A. Kim, P.T. Krein, Hot spotting and second breakdown effects on reverse I-V characteristics for mono-crystalline Si Photovoltaics, in *2013 IEEE Energy Conversion Congress and Exposition*, 2013, pp. 1007–1014
5. R. Alik, A. Jusoh, An enhanced P&O checking algorithm MPPT for high tracking efficiency of partially shaded PV module. *Solar Energy* **163**, 570–580, 2018/03/15/ (2018)
6. S. Motahhir, A. Chalh, A. El Ghzizal, A. Derouich, Development of a low-cost PV system using an improved INC algorithm and a PV panel Proteus model. *J. Clean. Prod.* **204**, 355–365, 2018/12/10/ (2018)
7. O. Guenounou, B. Dahhou, F. Chabour, Adaptive fuzzy controller based MPPT for photovoltaic systems. *Energy Convers. Manag.* **78**, 843–850, 2014/02/01/ (2014)
8. M. Mao, L. Cui, Q. Zhang, K. Guo, L. Zhou, H. Huang, Classification and summarization of solar photovoltaic MPPT techniques: a review based on traditional and intelligent control strategies. *Energy Rep.* **6**, 1312–1327, 2020/11/01/ (2020)
9. M. Seyedmahmoudian, B. Horan, T.K. Soon, R. Rahmani, A.M. Than Oo, S. Mekhilef, et al., State of the art artificial intelligence-based MPPT techniques for mitigating partial shading effects on PV systems—a review. *Renew. Sustain. Energy Rev.* **64**, 435–455, 2016/10/01/ (2016)
10. K. Ishaque, Z. Salam, M. Amjad, S. Mekhilef, An improved particle swarm optimization (PSO)-based MPPT for PV with reduced steady-state oscillation. *IEEE Trans. Power Electron.* **27**, 3627–3638 (2012)
11. C. Huang, Z. Zhang, L. Wang, Z. Song, and H. Long, A novel global maximum power point tracking method for PV system using Jaya algorithm, in *2017 IEEE Conference on Energy Internet and Energy System Integration (EI2)*, 2017, pp. 1–5
12. M.S. Sheik, Multiple step size perturb and observe maximum power point tracking algorithm with zero oscillation for solar PV applications, in *2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)*, 2018, pp. 1–5
13. K.S. Tey, S. Mekhilef, Modified incremental conductance algorithm for photovoltaic system under partial shading conditions and load variation. *IEEE Trans. Industr. Electron.* **61**, 5384–5392 (2014)
14. S. Motahhir, A. El Ghzizal, S. Sebti, A. Derouich, Modeling of photovoltaic system with modified incremental conductance algorithm for fast changes of irradiance. *International Journal of Photoenergy* 2018, 3286479, 2018/03/13 (2018)

15. X. Li, H. Wen, Y. Hu, and L. Jiang, A novel beta parameter based fuzzy-logic controller for photovoltaic MPPT application. *Renew. Energy*, **130**, 416–427, 2019/01/01/ (2019)
16. S. Farajdadian, S.M.H. Hosseini, Optimization of fuzzy-based MPPT controller via meta-heuristic techniques for stand-alone PV systems. *Int. J. Hydrogen Energy* **44**, 25457–25472, 2019/10/04/ (2019)
17. H. Hamdi, C. Ben Regaya, A. Zaafour, Real-time study of a photovoltaic system with boost converter using the PSO-RBF neural network algorithms in a MyRio controller. *Solar Energy* **183**, 1–16, 2019/05/01/ (2019)
18. K.S. Tey, S. Mekhilef, M. Seyedmahmoudian, B. Horan, A.T. Oo, A. Stojcevski, Improved differential evolution-Based MPPT algorithm using SEPIC for PV systems under partial shading conditions and load variation. *IEEE Trans. Industr. Inf.* **14**, 4322–4333 (2018)
19. I. Shams, S. Mekhilef, K. S. Tey, Improved team game optimization algorithm based solar MPPT with fast convergence speed and fast response to load variations. *IEEE Trans. Ind. Electron.*, 1 (2020). <https://doi.org/10.1109/TIE.2020.3001798>
20. I. Shams, S. Mekhilef, K. S. Tey, Maximum power point tracking using modified Butterfly optimization algorithm for partial shading, uniform shading and fast varying load conditions. *IEEE Trans. Power Electron.* **36**(5), 5569–5581 (2021). <https://doi.org/10.1109/TPEL.2020.3029607>
21. D.S. Pillai, J.P. Ram, A.M.Y.M. Ghias, M.A. Mahmud, N. Rajasekar, An accurate, shade detection-based hybrid maximum power point tracking approach for PV systems. *IEEE Trans. Power Electron.* **35**, 6594–6608 (2020)

Protection Scheme Based on k-Nearest Neighbour Algorithm for Fault Detection Classification and Section Identification in DC Microgrid



Shankarshan Prasad Tiwari, Ebha Koley, and Subhojit Ghosh

Abstract The rapid increase in the power requirement along with the stress on reducing the dependence on fossil fuels has propelled the integration of green sources of energy such as PV and wind energy. The increase in the use of DC power-based devices along with the ease of integrating renewable sources has led to significant interest in DC microgrids. However, the absence of zero crossing, bi-directional current flow and the dependence of the fault current magnitude on the operating modes (islanded and grid connected) poses significant challenges in designing a reliable protection scheme for DC microgrid. In this regard, a k-nearest neighbour (kNN)-based scheme has been proposed in the present work to perform the task of mode detection, fault detection/classification and section identification in DC microgrid. The algorithm does not involve a module for extracting features from the post-fault time waveforms, thereby leading to faster execution of the protection tasks. The scheme has been extensively validated for varying fault scenarios in terms of accuracy and computational cost.

Keywords DC microgrid · kNN · Distributed energy resources · DC microgrid protection

1 Introduction

The growing power requirements during past few years and the anticipated depletion of non-renewable sources have encouraged the adoption of green and alternate sources of energy [1]. In recent times, considerable investigation and development have been carried out to amalgamate renewable distributed energy resources (DERs) such as photovoltaic (PV) generator and wind turbine generator into existing power

S. P. Tiwari · E. Koley (✉) · S. Ghosh
Department of Electrical Engineering, N.I.T Raipur, Raipur, C.G, India
e-mail: ekoley.ele@nitrr.ac.in

S. Ghosh
e-mail: sghosh.ele@nitrr.ac.in

distribution networks [2]. The same has led to an increased interest among researchers towards the utilization of renewable energy resources in the low and medium power distribution network [3, 4]. The integration of distributed energy resources in the distribution network to provide reliable power requirement has contributed significantly towards the establishment of a single controllable entity referred to as microgrid [5, 6]. The microgrid consists of distributed energy resources (DERs), energy storage systems and electrical load at the consumer end. The microgrid pertains to small-rating power grid structure [7–9], which can be classified as alternating current (AC), direct current (DC) or hybrid depending on the type of connection. The drastic increase in the use of electronics devices such as mobile phones, lighting emitting diodes (LEDs), grinders, savers and other DC power fed equipment has led to the requirement of DC microgrid. DC microgrids have found application in different domains like small-scale power generation, backup of energy storages, data centres, marine and critical loads. In DC system, the power conversion stages are reduced as compared to the AC system, which results in reduced loss, high efficiency and increased reliability, so the system has got several benefits over the AC system [10]. However, DC microgrids are quite sensitive to fault and variation in the loading condition [11]. Also, the varying level of fault current during grid-connected and islanded mode makes the task of fault detection and classification challenging in DC microgrids since the threshold setting differs significantly between the operating mode [10]. A number of schemes have been documented in the literature regarding the fault detection and isolation of the DERs in post-fault scenarios. The notable schemes proposed for DC microgrid protection include superimposed current-based unit protection scheme [12], centralize unit-based protection scheme [13], wavelet transform-based efficient protection scheme for low-voltage DC microgrid [14], overcurrent protection [15] and hierarchical protection scheme with local differential protection [16]. In spite of their effectiveness in protecting DC microgrid for different fault types, the reported techniques are not able to execute the multiple tasks of fault detection/classification and identification of faulty section for different operating and loading conditions. These techniques are not able to detect the dynamic behaviour of distributed energy resources. In this context, a computationally simple algorithm based on kNN has been proposed for detecting the faults, determining the faulty phases and faulty section identification. The scheme has been extensively validated for different faults and operating scenarios of DC microgrid. The rest of the article is organized as follows: the test microgrid model is depicted in Sect. 2. The outline of kNN-based scheme is dealt in Sect. 3. Section 4 describes the development of the protection scheme. Section 5 is dedicated for result analysis while Sect. 6 for conclusions.

2 Test DC Microgrid Model Under Study

A 350 V test DC ring microgrid model is depicted in Fig. 1 [2]. The test microgrid model and the proposed protection scheme are simulated and implemented in

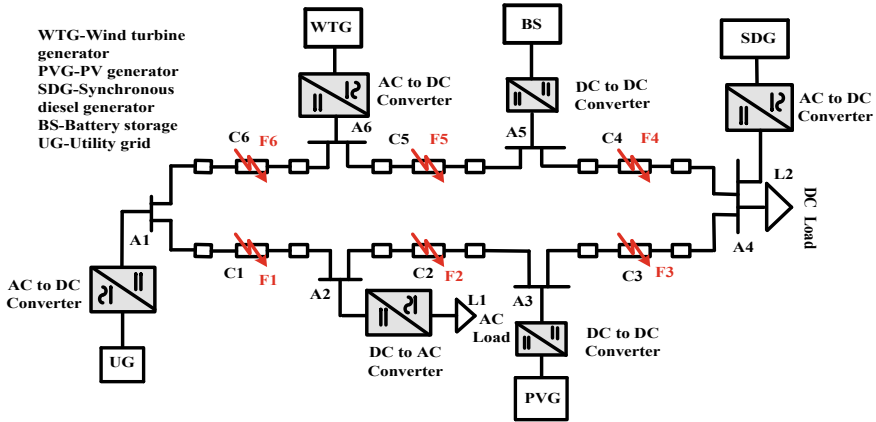


Fig. 1 Schematic representation of test DC microgrid model

MATLAB/Simulink. The system consists of three DERs, namely solar, synchronous diesel generator and wind turbine generator with the microgrid divided into six sections C1, C2, C3, C4, C5 and C6. The microgrid extends over the length of 6 kilometres with each section comprising of one kilometre.

3 Outline of KNN-Based Protection Scheme

The kNN is simple and supervised machine learning-based non-parametric classifier, which works on the principle of pattern reorganization. The algorithm initiates with finding the k-nearest neighbours of a given feature vector in the feature space [17, 18]. The extent of nearness is quantified in terms of the Euclidean distance. The algorithm is based on the assumption that similar attributes appear closely in the feature space.

It is desirable to prefer odd value of neighbour to avoid the situations of tie condition during the prediction of given class. The given class is decided by neighbouring points among all the inputs. Suppose if test cases are Y for class C_i , then probability of cases is given by

$$kNN(Y) = \max_p(C_i, Y) \tag{1}$$

where p refers to the probability to which the vector Y belongs to C_i .

For performing the task of classification, weights are assigned in such a manner so that the nearest neighbours contribute more weights as compared to the farther ones according to the following equation [15]. The Euclidean distance of Y to a data point in the p -dimensional feature space is calculated as:

$$d(y, t) = \sqrt{(y_1 - t_1)^2 + (y_2 - t_2)^2 + \dots + (y_p - t_p)^2} = \sqrt{\sum_{i=1}^p (y_i - t_i)^2} \quad (2)$$

In the present work, the protection tasks have been formulated as a classification problem and further solved using kNN classifier. The kNN performs the task of fault detection/classification (FD/C) and faulty section identification (FS/I) under both modes of operation, i.e. grid-connected (GC) and islanded mode (IM). To perform the fault detection task, two targets, namely fault and no fault have been assigned, while for fault classification, the classifier assigns the output pole to ground (PG) and pole to pole (PP).

4 Development of Proposed Protection Scheme

In the proposed protection scheme, the post-fault samples of voltage and current are retrieved from the relaying bus A1. The outline of the proposed kNN-based classifier is illustrated in Fig. 2. Diverse scenarios involving variation in fault operating parameters like fault resistance, fault location and load have been considered to train the kNN classifier. The mode detection task is carried out by kNN-1 with two output classes, i.e. GC and IM mode. Following the detection of the operating mode, the development of proposed scheme for fault detection/classification and section identification

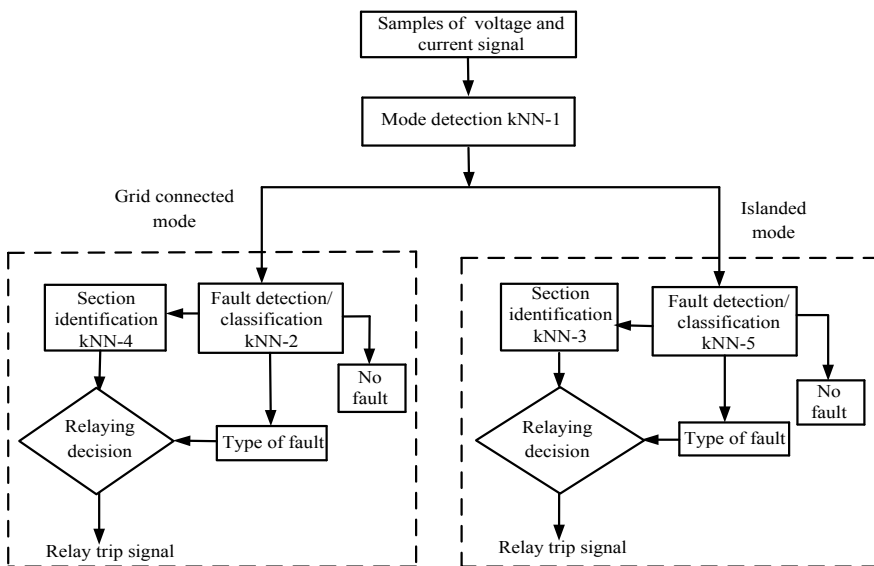


Fig. 2 Protection scheme of kNN-based classifier

is carried by the corresponding kNN modules. It is to be noted that inaccurate detection of the operating mode will lead to inaccuracy in fault detection/classification and section identification.

4.1 Fault Detection/Classification

In this subsection, the architecture of FD/C modules are discussed to design the proposed algorithm. The task of FD/C is performed by kNN-2 during the grid-connected mode while kNN-5 during the islanded mode of operation, respectively. To perform the assigned task of FD/C, sampled voltage-current signals have been considered as the input data to be fed to the kNN. The respective kNN modules provide the output regarding the faulty or healthy condition with type of fault among the probable fault types, i.e. pole to ground and pole to pole fault.

4.2 Faulty Section Identification

This subsection is dedicated to describe the section identification module. After mode detection and FD/C, fault section identification (FS/I) is also an important task. Therefore, for section identification, the modules kNN-4 and kNN-3 are developed under GC and IM mode, respectively. As outlined earlier, in the proposed protection model, six sections have been considered for identifying the faulty section.

5 Result Analysis

A dataset of 726 pole to ground (PG), 66 pole to pole (PP) fault and 100 NF (no fault) cases in each mode has been considered for validating the proposed kNN-based protection scheme. The performance of proposed algorithm under both modes is dealt below.

5.1 Mode Detection

For mode detection module, a total number of 480 test cases in each mode have been considered. The kNN-1 module has been tested with two classes, namely 'GC' and 'IM'. After successful testing, it has been examined that the accuracy of mode detector is 100% during the grid-connected mode and islanded mode. The higher level of accuracy indicates the effectiveness of mode detection module kNN-1.

5.2 Fault Detection/Classification

The performance of FD/C under both modes is summarized in Table 1. To analyse the appropriateness of fault detector/classifier, a total number of 676 cases with 100 no-fault cases have been considered. The high testing accuracy of fault detector/classifier reflects its efficacy in accurately detecting and classifying the fault irrespective of the operating scenario.

To realize the fault condition during an abnormal condition of power distribution network, a PG fault in section C1 is simulated at $t = 0.4$ s under both modes, and the variation in magnitude of voltage and current is depicted in Fig. 3. The waveform in Fig. 3a, b indicates the variation in the magnitude of voltage and current during grid-connected mode, while the post-fault dynamics are shown in Fig. 3c, d.

Table 1 Testing accuracy (%) of fault detection/classification

Type of fault	Accuracy of fault detection/classification (%)		Overall accuracy (%)
	Grid-connected mode	Islanded mode	
Pole to ground fault	100	98.41	99.20
Pole to pole fault	98.93	97.81	98.37

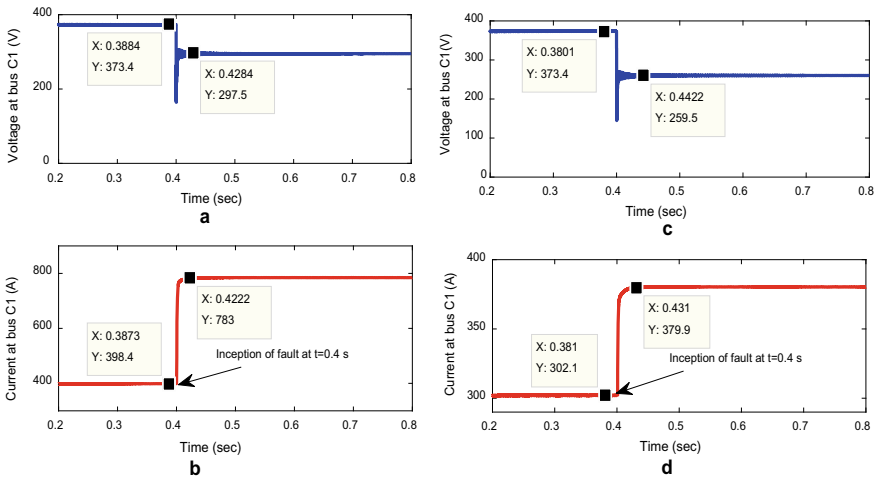


Fig. 3 Variation of DC voltage and current during pole to ground fault in section C1 for grid-connected (a, b) and islanded (c, and d) mode

Table 2 Testing accuracy of faulty section identification

Section	Mode of operation (accuracy in percentage %)		Overall accuracy percentage (%)
	Grid-connected mode	Islanded mode	
C1	99.27	98.39	98.83
C2	98.69	98.28	98.48
C3	98.66	97.31	97.98
C4	99.15	98.11	98.63
C5	98.29	97.11	97.70
C6	98.23	97.38	97.80

5.3 Faulty Section Identification

In this subsection, the performance of kNN-based module for section identification has been analysed. The outputs of kNN-3 and kNN-4 modules are summarized in Table 2. A total of 318 test cases for each mode have been considered to test the classifiers. The test results reflect the accuracy of the proposed scheme in identifying the faulty section with increased robustness to the operating fault scenarios.

6 Conclusion

A protection scheme based on k-nearest neighbour algorithm has been proposed for DC microgrid. The challenges pertaining to the amalgamation of renewable sources in the power distribution network have been addressed by formulating the protection tasks as a classification problem and further solving it using kNN algorithm. The dependence of the network dynamics on the operating mode has been dealt by designing a dedicated kNN module for mode detection. Further to mode detection, separate modules have been designed for fault detection/classification and section identification. The proposed scheme has been extensively validated for various faults and no-fault scenarios under diverse conditions. The test results confirm the robustness and reliability of protection modules.

References

1. M. Manohar, E. Koley, SVM based protection scheme for microgrid, in *International Conference on Intelligent Computing, Instrumentation and Control Technologies*, pp. 429–432 (2017)
2. R. Mohanty, A.K. Pradhan, Protection of smart DC microgrid with ring configuration using parameter estimation approach. *IEEE Trans. Smart Grid* **9**, 6328–6337 (2017)

3. M. Manohar, E. Koley, S. Ghosh. A reliable fault detection and classification scheme based on wavelet transform and ensemble of SVM for microgrid protection, in *International Conference on Applied and Theoretical Computing and Communication Technology*, pp. 24–28 (2017)
4. F. Firuzabad, Mahmud, R. Iravani, F. Aminifar, N. Hatziargyriou, M. Lehtonen. Guest editorial special section on microgrids. *IEEE Trans. Smart Grid* 1857–1859 (2012)
5. S.A. Hosseini, A. A. Hossein, Seyed Hossein, S.H.H. Sadeghi, F. Razavi, A. Nasiri, An overview of microgrid protection methods and the factors involved. *Renew. Sustain. Energy Rev.* **64**, 174–186 (2016)
6. R.H. Lasseter, Microgrid, in *IEEE Power and Energy Society Winter Meeting*, vol. 1, pp. 305–308 (2002)
7. S. Mirsaedi, D. Xinzhou, D. Mat Said, Towards hybrid AC/DC microgrids: critical analysis and classification of protection strategies. *Renew. Sustain. Energy Rev.* **90**, 97–103 (2018)
8. M. Manohar, E. Koley, S. Ghosh, Enhancing the reliability of protection scheme for PV integrated microgrid by discriminating between array faults and symmetrical line faults using sparse auto encoder. *IET Renew. Power Gener.* **13**(2), 308–317 (2018)
9. M. Manohar, E. Koley, S. Ghosh, Stochastic weather modeling-based protection scheme for hybrid PV–wind system with immunity against solar irradiance and wind speed. *IEEE Syst. J.* **14**(3), 3430–3439 (2020)
10. N. Bayati, A. Hajizadeh, M. Soltani, Protection in DC microgrids: a comparative review. *IET Smart Grid* **1**(3), 66–75 (2018)
11. A. Meghwani, S.C. Srivastava, S. Chakrabarti, A non-unit protection scheme for DC microgrid based on local measurements. *IEEE Trans. Power Deliv.* **32**(1), 172–181 (2016)
12. R. Mohanty, A.K. Pradhan, A superimposed current based unit protection scheme for DC microgrid. *IEEE Trans. Smart Grid* **9**(4), 3917–3919 (2018)
13. M. Monadi, C. Gavriluta, A. Luna J.I., Candela, P. Rodriguez, Centralized protection strategy for medium voltage DC microgrids. *IEEE Trans. Power Deliv.* **32**(1), 430–440 (2016)
14. S. Som, S.R. Samantaray, Efficient protection scheme for low-voltage DC micro-grid. *IET Gener. Trans. Distri.* **12**(13), 3322–3329 (2018)
15. D.Salomonsson, L. Soder, A. Sannino, Protection of low-voltage dc microgrids. *IEEE Trans. Power Deliv.* **24**(3), 1045–1053 (2009)
16. L. Che, M.E. Khodayar, M. Shahidehpour, Adaptive protection system for microgrids: protection practices of a functional microgrid system. *IEEE Electr. Mag.* **2**(1), 66–80 (2014)
17. Y. Xiao-Peng, Y. Xiao-Gao. Novel text classification based on k-nearest neighbour, in *International Conference on Machine Learning and Cybernetics*, vol. 6, pp. 3425–3430 (2007)
18. S.K. Shukla, Ebha Koley, Fault detection and classification in six-phase transmission system using k-nearest neighbour algorithm, in *International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT)*, pp. 542–546. 2017

Solar PV-Fed Grid-Interactive PMSM-Driven Water Pumping System with Switched-LC Converter



Rasna Rajan, C. M. Nirmal Mukundan, M. K. P. Muhammad Ramees, and P. Jayaprakash

Abstract Solar-based water pumping has enhanced the feasibility of pumping application since water pumping is an essential operation in agriculture. A solar photovoltaic (PV) water pumping system having two stages with a grid interaction is presented in this article. A switched-LC high-gain converter with a combination of switched capacitor and switched inductor is used at first stage as boost converter which is fed with a PV input and a single-phase grid for an effectual power transfer between them to feed a centrifugal pump connected with a permanent magnet synchronous motor (PMSM) drive through a three-phase voltage source inverter (VSI). An uninterrupted pumping operation is provided to the consumer with grid integration at any insolation range. When no pumping is required, the PV power feeds the utility grid. To control converter duty ratio and to extract maximum energy from PV panel, the maximum power point tracking is implemented using incremental conductance (INC) algorithm. A field-oriented control technique is carried out to drive the motor to rotate the pump. The simulation of the system is done in MATLAB, and results are developed.

Keywords PV panel · Incremental conductance algorithm · Utility grid · Field-oriented control · Permanent magnet synchronous motor

1 Introduction

Renewable energy source is clean, pollution-free energy sources of nature that encouraged the dependency of humans on it. Since the population growth rate is increasing massively, the electricity demand is also growing which is enhancing the technology on different renewable sources [1]. The abundance of solar energy has developed the technology on utilization of solar PV-based application since PV cells are simple in design, low cost and maintenance and, moreover, durable in nature. Presently, the domestic and irrigation areas are utilizing the PV energy for

R. Rajan (✉) · C. M. Nirmal Mukundan · M. K. P. Muhammad Ramees · P. Jayaprakash
Department of Electrical and Electronics Engineering, Government College of Engineering
Kannur, Kannur, Kerala, India

the pumping application to have a sustainable water pumping operation. For operating the pump, AC electric drives are preferred than DC. Even though the popularity of induction motor (IM) drives in pumping application is high due its robust nature and low maintenance cost, the high reactive power demand made the drive relatively inefficient for the application. The advantages of permanent magnet synchronous motor (PMSM) drive like small size, high efficiency, high power density and low rotor inertia made it more applicable drive for pumping than IM [2]. Field-oriented control (FOC) and direct torque control (DTC) are two general vector control methods for electric motor drives. FOC is better for PMSM drive control than DTC since the presence of permanent magnets in rotor of the drive makes the FOC less sensitive to rotor resistance while it shows larger torque ripples toward DTC. Thus, FOC-based PMSM drive has good steady state and dynamic performance. A step-up converter boosts the level of the voltage. Among isolated and non-isolated types of converters, the absence of isolation transformers makes the non-isolated type preferred over isolated type. The variation of duty cycle (D) widely between 0 and 1 rises the device stress. As a solution to this, Z-source converter [3] was developed which expressed limitations like high starting current and high input current. Coupled inductor-based converters in [4] generated spikes in the circuit due to severe leakage inductance. This dropped the gain of the voltage largely. This work presents a two-switch high-gain step-up converter with switched inductor/capacitor with a separate interconnection between input and output through a diode. The device voltage stress and diode conduction losses are reduced, providing better gain than networks [5]. To overcome the problems of intermittent power supply in standalone mode leading to unreliable water pumping, a grid can be integrated through a bidirectional topology capable of feeding the pump and storing extra power from PV and hence provide a continuous pumping operation.

In this paper, a PMSM-driven utility grid integrated WPS is presented. Instead of a conventional boost converter, a switched-LC high-gain converter is utilized in this work with an INC algorithm to achieve MPPT of PV array. FOC using hysteresis current control is used for motor speed control. A unit vector template (UVT) theory is employed for bidirectional power flow, with which the system focuses to maintain a constant rate of power flow to achieve an undisturbed pumping at any level of solar insolation. This system evades any components for switching power source between grid and PV panel and vice versa [6, 7].

The paper is ordered as follows: Sect. 2 discusses system architecture, Sect. 3 mentions the control schemes used, Sect. 4 shows the system design, and Sect. 5 presents the simulation results, Sect. 6 concludes the chapter.

2 System Architecture

Figure 1 depicts the system diagram. This comprises a PV-fed high-gain switched-LC converter with MPPT control maintaining a constant DC link voltage [8, 9]. The PMSM coupled with the pump is driven by a three-phase voltage source inverter

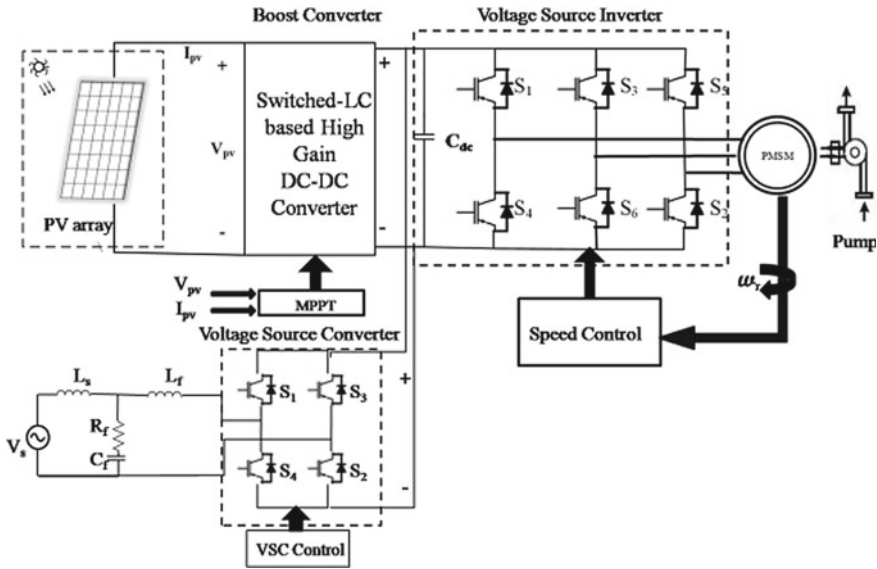


Fig. 1 System block diagram representation

(VSI). The same DC link is connected with a utility grid through a single-phase voltage source converter (VSC). The system operates in standalone mode and grid-connected mode. In grid-connected mode, there are three cases based on feeding operation—first case: grid to pump; second case: both PV and grid to pump; and third case: PV to grid.

2.1 Switched-LC High-Gain Converter

Figure 2 shows the converter presented has a switched inductor and switched capacitor combination. The gain of the presented topology is comparatively higher gain than other converters performed in low duty cycle ranges [10].

There are two modes of converter operation: continuous current mode (CCM) and discontinuous current mode (DCM). To have better efficiency and lower inductor current ripple, this paper considers only CCM. Capacitors are large enough to have same voltages across it during charging and discharging mode. There are two modes suitable for this converter in CCM: charging mode, where both switches are ON and the inductors L_1 and L_2 charge. In this mode, the power to load is fed from the capacitor voltages V_{C2} , V_{C3} and source voltage V_{in} . In discharging mode, both switches are in OFF state and the inductors L_1 and L_2 discharge. The capacitor C_0 maintains a constant load voltage V_0 in this mode.

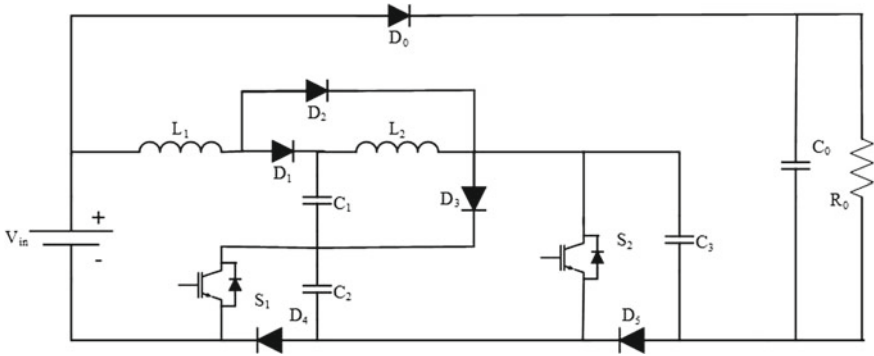


Fig. 2 Switched-LC high-gain converter

3 Control Schemes

The control schemes for the whole system include an INC algorithm for DC link voltage control of the converter. PMSM speed is controlled using FOC based on hysteresis-band current controlled VSI, and unit vector template (UVT) is used to control VSC connected with grid to make the power to flow bidirectionally. This ensures a continuous power input for the pump even when the insolation level falls down largely.

4 Design of the System

This section details the system design. Sections 4.1 and 4.2 give the PV array and converter design. Section 4.3 shows the design of pump load used in this system.

4.1 PV Array

1 kW PV array is designed for a 1 hp, 3-phase, 220 V PMSM drive with a rated speed of 1500 rpm. The parameters estimated at STC (1000 W/m², 25 °C) and ratings of designed PV module are indexed in Table 1.

4.2 Design of Switched-LC High-Gain Converter

The converter voltage gain (G) is obtained by Eq. (1), where D is the duty ratio.

Table 1 Specification of solar PV module

Parameter	Value
MPP power (P_{mp})	179.1 W
Open circuit voltage (V_{oc})	31.7 V
MPP voltage (V_{mp})	26.9 V
Short circuit current (I_{sc})	7.77 A
MPP current (I_{mp})	21.8 A

Table 2 Specification of converter

Parameter	Value
Frequency, f_{sw}	10 kHz
Inductance L_1, L_2	2 mH, 0.55 mH
Capacitance C_1, C_2, C_3	10 μ F, 10 μ F, 10 μ F
Capacitance C_0	47 μ F

$$G = \frac{V_0}{V_{in}} = \frac{3 - 3D + D^2}{1 - 3D + D^2} \tag{1}$$

Different parameters of the converters are designed based on [6], and the ratings are indexed in Table 2.

4.3 Design of Pump

A rotating pump is coupled to drive. The speed and torque characteristics of load show that load torque (T_1) is proportional to square of motor speed (ω_r).

$$K_p = \frac{T_l}{\omega_r^2} = \frac{4.75}{157^2} = \frac{0.000192 \text{ Nm}}{(\text{rad/s})^2} \tag{2}$$

where is K_p the pump constant and $\omega_r = (N_r * \pi)/30$. N_r is the rated motor speed.

5 Simulation Results

At varying of operating conditions, the starting, stable state and dynamic characteristics are validated during changes in insolation and temperature changes. All possible conditions are included—operation of PV array at full rating, load fed by grid alone, both grid and PV feeding pump and grid fed by PV array.

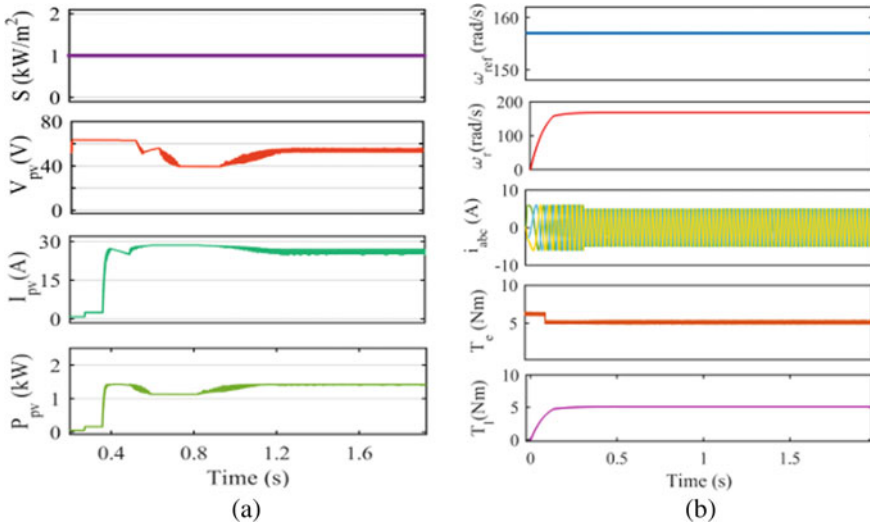


Fig. 3 Output waveforms of **a** PV panel, **b** PMSM at starting and stable-state period when pump fed by PV

5.1 Performance at Starting and Stable State

The performance of the system at starting and stable state during all three cases in grid-connected mode mentioned in Sect. 2 is analyzed.

- (a) **PV panel feeding pump:** The performance of the PV and PMSM during starting and reaching a steady-state period are presented in Fig. 3a, b. The condition for full capacity pump operation is during insolation of 1000 W/m^2 .
- (b) **Pump fed by Grid:** Fig. 4a illustrates the characteristics of the grid where drive draws a grid current of 5 A out of phase with the grid voltage. The performance of the drive at starting and reaching a steady-state period is presented in Fig. 4b.
- (c) **PV panel feeding grid:** The PV array characteristics are depicted in Fig. 4c, and grid performance is presented in Fig. 4d. The current drawn by the grid is in phase with grid voltage. The DC link voltage around 400 V is maintained.

The indices for different parameters are ‘S’ for insolation range, ‘ P_{pv} , V_{pv} , I_{pv} ’ for PV power, voltage and current, ‘ ω_{ref} , ω_r , i_{abc} , T_e , T_l ’ for PMSM parameters.

5.2 Dynamic Performance During Irradiance Change

The dynamic behavior of the PV input, PMSM drive and the utility grid during insolation changes from 500 to 1000 W/m^2 at 1.4 s is presented in Fig. 5a–c. The

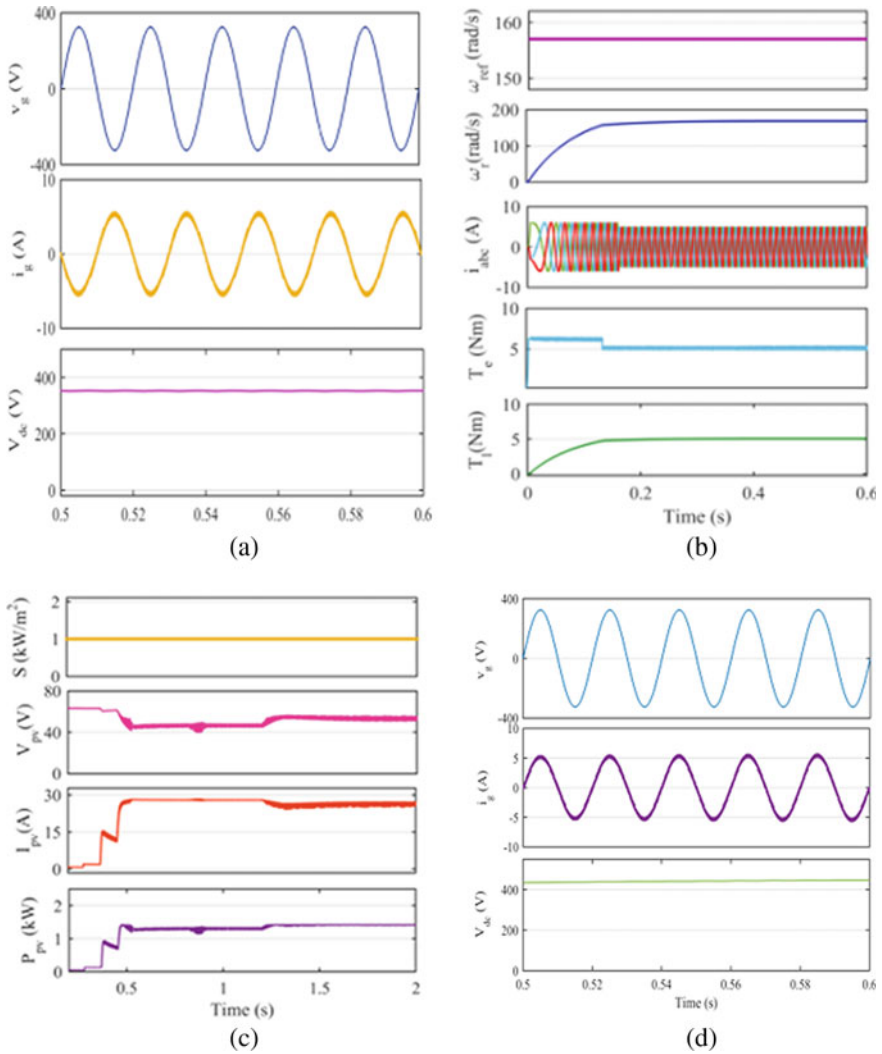


Fig. 4 Output waveforms at starting and stable- state period (i) when pump fed by grid-(a) grid (b) PMSM and (ii) when grid fed by PV-(c) PV panel (d) grid

drive is operating at the rated condition throughout the changes. The rated speed of motor of 157 rad/s is maintained. The voltage around 400 V at DC link is maintained.

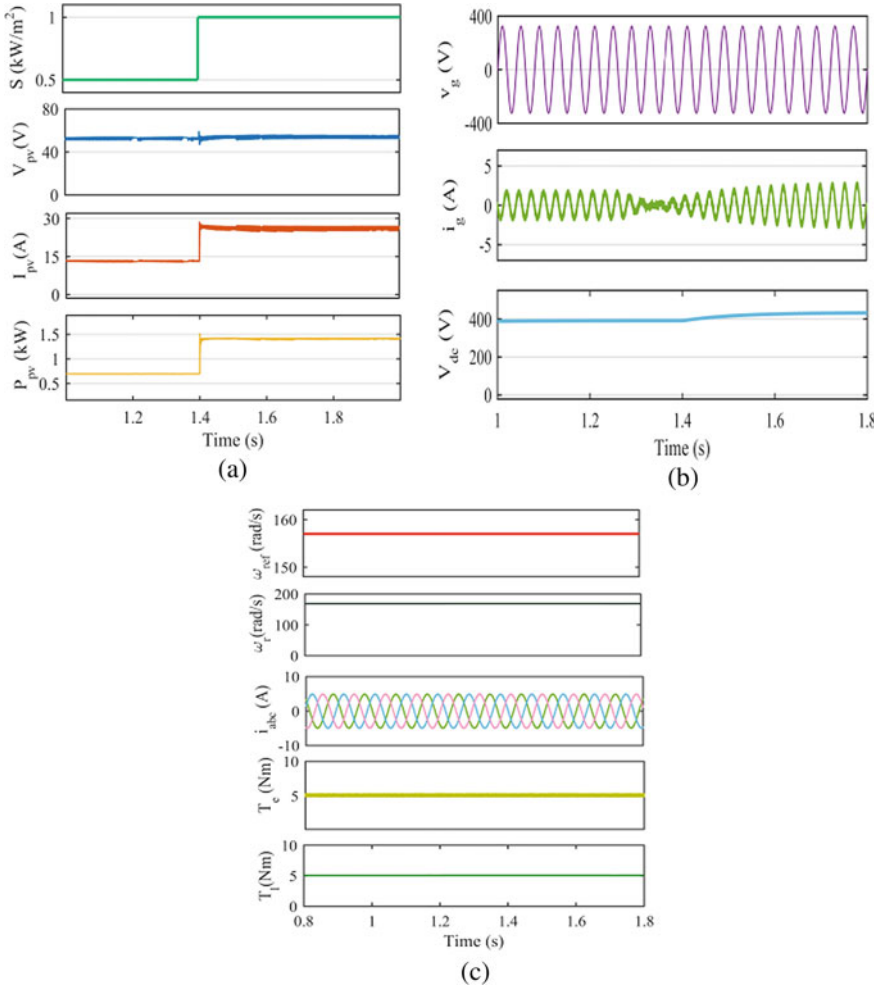


Fig. 5 Output waveforms of **a** PV array, **b** PMSM drive, **c** utility grid when insolation changes from 500 to 1000 W/m²

5.3 Power Quality

The total harmonic distortion of grid current (THD_i) attained when PV feeds the grid and when grid feeds the pump is presented in Fig. 6a, b. In both the cases, THD_i value is within 5%, which highly satisfies the IEEE-519 standard.

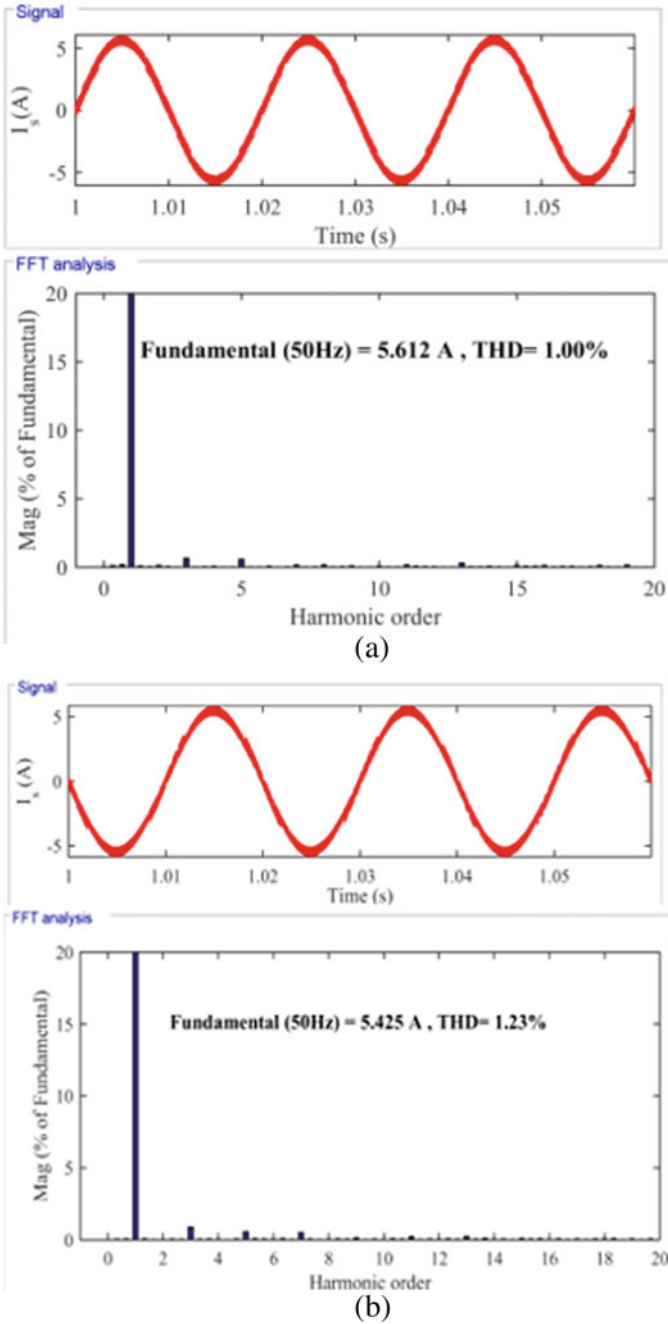


Fig. 6 Utility grid current harmonic spectra during a PV array feeds grid, b grid feeds pump

6 Conclusions

This paper has presented a grid-interactive PV array feeding PMSM-driven water pumping system. The boost converter used is a switched-LC-based high-gain topology. This converter has provided a better gain of voltage at low duty cycle, low voltage across switches, capacitors and diodes. An incremental conductance algorithm for the converter and control of speed of the drive by field-oriented control is the control operations used for the system. Control over bidirectional flow of power is executed using unit vector template for an uninterrupted pump utilization using a utility grid. The starting, stable state and dynamic performance of the drive are analyzed. The results achieved proved a satisfactory performance of the presented system and hence shown the fitness under all varying operating conditions. The system enhances a sustainable power management and a path of income with electricity delivery to the grid.

References

1. B. Singh and S. Murshid, A grid-interactive permanent-magnet synchronous motor-driven solar water-pumping system. *IEEE Trans. Power Electron.* **54**(5), 5549–5561 (2018)
2. T.R. Brinner, R.H. McCoy, T. Kopecky, Induction versus permanent-magnet motor for electric submersible pump field and laboratory comparisons. *IEEE Trans. Ind. Appl.* **50**(1), 174–181 (2014)
3. F.Z. Peng, X. Yuan, X. Fang, Z. Qian, Z. Qian, Z-source inverter for adjustable speed drives. *IEEE Trans. Power Electron. Lett.* **1**(2), 33–35 (2003)
4. R.N. Shaw, P. Walde, A. Ghosh, A new model to enhance the power and performances of 4×4 PV arrays with puzzle shade dispersion. *Int. J. Innov. Technol. Explor. Eng.* **8**(12) (2019)
5. Y.P. Siwakoti, F. Blaabjerg, V.P. Galigekere, A. Ayachit, M.K. Kazimierczuk, A-source impedance network. *IEEE Trans. Power Electron.* **31**(12), 8081–8087 (2016)
6. G. Zhang, B. Zhang, Z. Li, 3-z-network boost converter, in *Designing Impedance Networks Converters* (Springer, Berlin, Germany, 2018), pp. 55–82
7. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
8. D. Saravanan, A. Hasan, A. Singh, H. Mansoor, R.N. Shaw, Fault prediction of transformer using machine learning and DGA, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, 2020, pp. 1–5. <https://doi.org/10.1109/gucon48875.2020.9231086>
9. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON) held at Deenbandhu Chhotu Ram University of Science and Technology, SONPAT*. India on Feb 28–March 1, 2020
10. A. Kumar et al., Switched-LC based high gain converter with lower component count. *IEEE Trans. Ind. Appl.* **56**(3), 2816–2827, (2020)

Impact Analysis of Microgrid in Minimization of Distribution System Losses



Nawaz Ali Warsi, Anwar Shahzad Siddiqui, Sheeraz Kirmani, and Md Sarwar

Abstract Transmission and distribution systems are two important components of power system which are responsible for the availability of electricity to the end users. Its performance plays a crucial role in achieving the electricity market economics. However, the losses in transmission and distribution system are high in which the distribution system loss has a major share. This is due to high value resistance to reactance ratio, high current low voltage, etc. Therefore, in this paper, an attempt is made to minimize the distribution loss to enhance the performance of distribution system. The loss reduction is achieved by employing microgrid in the network for its different penetration level. The method is tested on IEEE 33-node radial distribution system, and the results are found very encouraging.

Keywords Microgrid · Distributed energy resource · Transmission and distribution loss · Network reconfiguration

1 Introduction

One of the major causes for deregulation in power sector was the inefficient operation of the power distribution system. The lack of resources and unavailability of investment in its infrastructure development were some of its reasons responsible for its crumble operation. Further, the frequent increase in load demand also adds to its woes. Therefore, the distribution utilities in India were under immense pressure to improve its performance during the early years of the decade between 1991 to 2000. The restructuring of power sector and World Bank investment in distribution system infrastructure development had proved a panacea for the reeling distribution systems in India. This increases the performance of the distribution utilities. However, with the exponential increase of load and competitive electricity market, the transmission

N. A. Warsi (✉) · A. S. Siddiqui · S. Kirmani · Md.Sarwar
Department of Electrical Engineering, Jamia Millia Islamia, New Delhi 110025, India

S. Kirmani
Department of Electrical Engineering, Aligarh Muslim University, Aligarh, U.P., India

and distribution networks are being forced frequently to operate under overloaded condition. In such scenario, the network losses become a major concern as it impacts the performance of the network. Apart from it, management of network losses plays a major role in achieving the economic benefits of the system. Therefore, its minimization is the key to maintain the reliable and efficient operation of the distribution network.

The total power received by the distribution system is a function of losses in transmission system. It is given by the difference in total power generation and transmission losses. Similarly, the power received by the load is the function of distribution losses. But losses in distribution system are very high as compared with transmission system due to high ratio of resistance to reactance values, high current and low voltage. Therefore, power loss must be minimized to improve the distribution system efficiency.

In the past few decades, a number of researchers have adopted different techniques to minimize the power loss in distribution system [1]. These techniques mainly include network reconfiguration, allocation of capacitor, distributed generations (DG) and DSTATCOM, etc. The network reconfiguration technique is feasible for low-voltage distribution systems while capacitor allocation is feasible for high-voltage distribution systems. Network reconfiguration method was first proposed in [2]. The authors had employed discrete branch and bound technique for achieving the best configuration of the network to achieve minimization of loss. However, the method considered some approximations in network reconfiguration which is overcome by Shirmohammadi et al. in [3]. A network reconfiguration method based on high-voltage distribution system (HVDS) concept is used in [4] to minimize the loss. The authors had reconfigured the network by bifurcating the agricultural and residential loads. To minimize the loss, the agricultural load is supplied with high voltage by connecting the small capacity high-voltage transformers at the load points. However, the method involves a high capital cost. The loss minimization by the optimal allocation of capacitor is reported in [5–7] while a methodology proposed in [8, 9] utilized a combination of network reconfiguration and capacitor placement techniques to minimize the loss in distribution system. Besides capacitor placement and network reconfiguration, allocation of DG and STATCOM is being used extensively for loss reduction. Such a method is presented in [10] wherein authors had developed an optimization model employing DG. Similarly, DSTATCOM is employed in [11] to minimize the loss. To maximize the benefit from the use of DG and STATCOM, various authors have utilized a combination of these two techniques to minimize the loss. Such a method is reported in [12, 13].

The main objective of this paper is to analyze the effect of implementing microgrid distribution network. The microgrid implementation is measured in terms of distribution loss reduction.

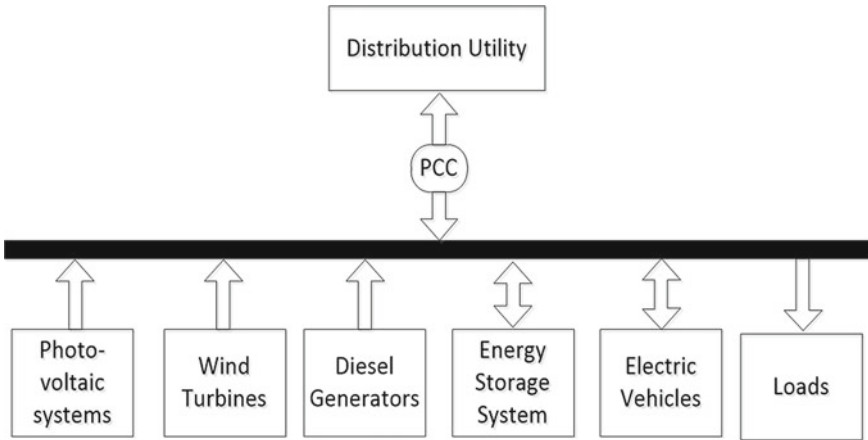


Fig. 1 Microgrid typical structure

2 Microgrid

To meet the exponentially increasing demand as well as to address the rising environmental concerns, the large-scale and centralized generating facilities are being gradually replaced by distributed energy resources (DERs) such as wind turbine, photovoltaic panel, fuel cell, diesel generator, energy storage devices like flywheel, supercapacitor and battery. However, for high-level penetration of DERs, it requires a robust monitoring and control over DERs which is very complex to achieve. Therefore, to ensure its better monitoring and control, the DERs are grouped in smaller numbers to form a microgrid. Thus, a microgrid can be seen as an interconnection of different small capacity energy sources (DERs) to meet the demand within electrically defined boundaries as shown in Fig. 1 [14]. The advantages of microgrid employing locally available distributed energy sources are apprehended through demand side management. It helps in achieving the objectives of distribution utilities.

3 Transmission and Distribution Line Losses

It is a fact that there is always a mismatch between the units of electrical energy generated by the power station and received by the consumers. This difference in electrical energy units is termed as transmission and distribution (T&D) loss. It measures the performance of T&D systems in terms of energy saving and is given by Eq. (1) [15]:

$$T \ \& \ D \ Losses = \left(1 - \frac{E_b}{E_i} \right) \times 100 \tag{1}$$

where E_i and E_b represent the total energy input to the system and total energy billed to consumers, respectively. The ratio (E_b/E_i) is known as billing efficiency of the distribution utilities.

Distribution loss constitutes a major portion of T&D losses. There are a number of reasons for high distribution losses, and some of which are as follows:

- Unequal load distribution.
- Lines overloading.
- Low voltages at consumer terminals.
- Abnormal operating conditions of distribution transformers.
- Large number of transformation stages.
- Poor power factor and inadequate compensation for reactive power.
- Poor maintenance, etc.

4 Results and Discussions

From the previous sections, it is evident that T&D loss is of utmost importance and needs greater concern for its minimization. It is also observed that the loss in the distribution network constitutes a major portion in T&D loss. Its minimization will have direct and major impact on the performance of the network in achieving the efficiency and economic benefits. Therefore, in this paper, to analyze the impact of the implementation of microgrid in minimizing the distribution on system losses, IEEE 33-node radial distribution system is considered [16]. A microgrid consisting of wind turbine, photovoltaic system and diesel generators is considered as DERs. One number of wind turbine at node 2, one number of photovoltaic system at node 4 and two numbers of dispatchable diesel generators at node 7 and node 8 in a 33-node system are considered and connected as shown in Fig. 2. The different ratings considered for these generators are shown in Table 1.

From Table 2, it is observed that the losses on the considered network reduce to 5.37% in case 1 when microgrid is employed in the network. For case 2 and case 3, the losses further reduced to 10.37% and 14.27%, respectively. Thus, it can be concluded

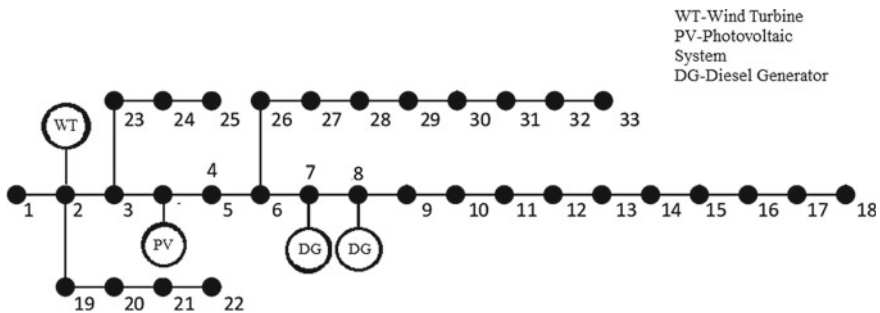


Fig. 2 IEEE 33-node radial distribution test system

Table 1 Ratings of generators

Generator types	Rating (kW)		
	Case 1	Case 2	Case 3
Wind turbine	25	25	25
Solar photovoltaic system	50	100	100
Diesel generator 1	50	100	150
Diesel generator 2	50	100	150

Table 2 System losses for different ratings of DERs of microgrid

Losses	Base case	Case 1	Case 2	Case 3
P_L (kW)	281.58	266.45	252.34	241.47
Q_L (kVAR)	187.95	177.89	168.53	161.04

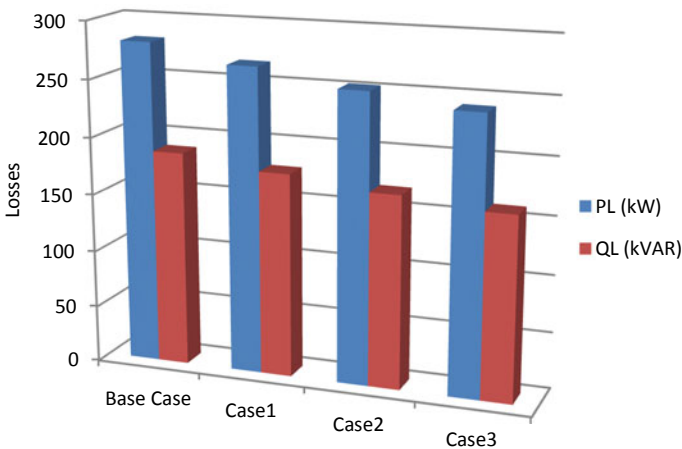


Fig. 3 System losses with microgrid

that the overall system losses reduced with the penetration of different levels of generations of microgrid generators. Hence, the efficiency of the distribution system is enhanced with the implementation of microgrid in the network. The reduction in system losses for each case is also illustrated in Fig. 3.

5 Conclusions

In this work, the impact of deployment of microgrid in distribution system is presented. The impact is measured in terms of reduction of system losses which increase the efficiency of the distribution system. A broad framework is presented for

benefit analysis of the microgrid-based generation for its different ratings. The effect is analyzed on IEEE 33-node radial distribution system. The results extracted from considered system show significant reduction of system losses. It is also observed that for the increased penetration level of microgrid generations, there is a significant reduction in system losses.

References

1. L. Ramesh, S.P. Chowdhry, S. Chowdhry, A.A. Natarajan, C.T. Gaunt, Minimization of power loss in distribution networks by different techniques. *Int. J. Electr. Electron. Eng.* **3**(9), 521–527 (2009)
2. A. Merlin, Search for a minimal-loss operating spanning tree configuration for an urban power distribution system, in *5th Power Systems Computations Conference*, Cambridge, UK, pp. 1–18 (1975)
3. D. Shirmohammadi, H.W. Hong, Reconfiguration of electric distribution networks for resistive losses reduction. *IEEE Trans. Power Delivery* **4**, 1402–1498 (1989)
4. M. Sarwar, Z.A. Jaffery, A.S. Siddiqui, I.A. Quadri, Techno-economic feasibility of HVDS concept for distribution feeder power loss minimisation, in *5th IEEE India International Conference on Power Electronics (IICPE)*, IEEE, New Delhi, India, pp. 1–4 (2012)
5. A. Mendes, P.M. Franca, C. Lyra, C. Pissarra, C. Cavellucci, Capacitor placement in large-sized radial distribution networks. *IEE Proc. Gener., Transm. Distrib.* **152**(4):496–502 (2005)
6. S.K. Injeti, V.K. Thunuguntla, M. Shareef, Optimal allocation of capacitor banks in radial distribution systems for minimization of real power loss and maximization of network savings using bio-inspired optimization algorithms. *Int. J. Electr. Power Energy Syst.* **69**, 441–455 (2015)
7. A.A. El-Fergany, Optimal capacitor allocations using evolutionary algorithms. *IET Gener. Transm. Distrib.* **7**(6), 593–601 (2013)
8. H.R. Esmaeilian, R. Fadaeinedjad, Distribution system efficiency improvement using network reconfiguration and capacitor allocation. *Int. J. Electr. Power Energy Syst.* **64**, 457–468 (2015)
9. J.M. Home-Ortiz, R. Vargas, L.H. Macedo, R. Romero, Joint reconfiguration of feeders and allocation of capacitor banks in radial distribution systems considering voltage-dependent models. *Int. J. Electr. Power Energy Syst.* **107**, 298–310 (2019)
10. S.A. Taher, M.H. Karimi, Optimal reconfiguration and DG allocation in balanced and unbalanced distribution systems. *Ain Shams Eng. J.* **5**(3), 735–749 (2014)
11. R.S. Somalwar, S.G. Kadwane, R.N. Shaw, Frequency estimation by recursive least square in active islanding method for microgrid, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, 2020, pp. 116–123. <https://doi.org/10.1109/gucon48875.2020.9231206>
12. S. Sannigrahi, P. Acharjee, Maximization of system benefits with the optimal placement of DG and DSTATCOM considering load variations. *Procedia Comput. Sci.* **143**, 694–701 (2018)
13. T. Yuvaraj, K. Ravi, Multi-objective simultaneous DG and DSTATCOM allocation in radial distribution networks using cuckoo searching algorithm. *Alexandria Eng. J.* **57**(4), 2729–2742 (2018)
14. E. Hossain, E. Kabalci, R. Bayindir, R. Perez, A comprehensive study on microgrid technology. *Int. J. Renew. Energy Res.* **4**(4), 1094–1107 (2014)
15. N.A. Warsi, A.S. Siddiqui, S. Kirmani, M. Sarwar, Impact assessment of microgrid in smart cities: Indian perspective. *Technol. Econ. Smart Grids Sustain. Energy* **4**(14), 1–16 (2019)
16. IEEE distribution planning working group report, Radial distribution test feeders. *IEEE Trans. Power Syst.* **6**(3), 975–985 (1991)

Performance Estimation and Analysis of 3D Trigate HOI FinFET Using Strained Channel for Reduced Area



Swagat Nanda and Rudra Sankar Dhar

Abstract In comparison with planar MOSFETs, multiple gate FETs such as trigate FinFETs are able to be effectively scaled down beyond 32 nm technology, with increasing device performance. But, below 20 nm, even the performance of trigate FinFETs degrade due to short channel effects. Therefore, the use of strained silicon in FinFET devices is a competent method to increase the drain current. This paper describes the implementation and simulation of 3D trigate FinFET with a gate length of 14 nm having a channel consisting of two strained silicon layers that sandwich SiGe between them. The device parameters such as I_{on} current, I_{off} current, threshold voltage, subthreshold slope, DIBL, and I_{on}/I_{off} current ratio are compared to another device having identical dimensions and doping but having only silicon as the channel material. The results showed improved device characteristics of the device using strained channel while consuming reduced chip area.

Keywords TG SOI FinFETs · Strained silicon · Heterostructure-on-insulator · Silvaco · Threshold voltage

1 Introduction

Due to the immense scaling of planar bulk MOSFETs to pack more devices in smaller area, the width of the depletion region formed by the source and drain of the MOSFET with the channel becomes comparable with the length of the channel in the sub-90 nm regime [1]. As the channel length diminishes, the channel barrier also decreases leading to threshold voltage roll off [2] and increased control of the channel by the applied drain voltage which affects the drain current [3] and becomes a major hindrance in further scaling of MOSFETs. Various literature indicate that

S. Nanda (✉) · R. S. Dhar
Department of Electronics and Communication Engineering, National Institute of Technology
Mizoram, Aizawl, India

R. S. Dhar
e-mail: rdhar@uwaterloo.ca

further scaling is possible by using either fully depleted silicon-on-insulator (FDSOI) MOSFET [4] or increasing the number of gates in the MOSFET structure [5]. In the FDSOI MOSFET, the buried oxide (BOX) layer improves the device performance by lowering the leakage currents resulting in higher sub-threshold slope. However, this increases the junction capacitance of the device. The multiple gate device approach led to the development of double gate (DG) and triple gate (TG) devices which shows promising results in the sub-90 nm technology. Among the different multi-gate device architectures, the 3D FinFET has emerged as the most efficient structure [6] to replace the planar MOSFET due to their fabrication simplicity.

As the FinFETs are scaled below 32 nm technology, the short channel effects resurface leading to lower drain currents. Therefore, newer research is needed to improve the drain currents of the devices at sub-32 nm technology nodes without degrading the short channel parameters further. As the drain current is proportional to the mobility of electrons and holes, strain technology can be used in the FinFETs to improve the drain current and reduce the short channel effects of sub-32 nm devices. Strain technology increases the channel carrier mobility, leading to a higher current flow in the channel, which can improve the device's performance [7]. Although research suggests that the application of strain is possible in different regions of the device, it is basically categorized as (i) global strain and (ii) local strain. A global strain is applied when the strain is directly used in the substrate (strained SOI) [8, 9]. The local strain technology may consist of stressed metal contacts and gates [10], stress memorization techniques [11, 12], stressed over-layers [13, 14], or embedded source-drain-stressor [15].

In [16], strained silicon region is generated by using only Si and only Ge strained layers on relaxed $\text{Si}_{1-x}\text{Ge}_x$ ($x \sim 0.7$) resulting in high mismatch between Si and $\text{Si}_{0.3}\text{Ge}_{0.7}$ causing high density dislocations, displaying higher degradation of electron mobility, while the mobility of holes remains unchanged. Strained silicon regions can also be developed by using strained silicon on relaxed $\text{Si}_{1-x}\text{Ge}_x$, or strained Si and strained $\text{Si}_{1-x}\text{Ge}_x$ on relaxed $\text{Si}_{1-y}\text{Ge}_y$ ($x > y$) as well as strained Si and strained Ge on $\text{Si}_{1-y}\text{Ge}_y$ FET devices [16–20]. It becomes beneficial to keep a lower percentage of germanium in the $\text{Si}_{1-x}\text{Ge}_x$ alloy since it provides better results than using pure Ge for the strained silicon channel FinFETs. In [21], a ~49% enhancement in drain current is achieved by employing a double-strained Si channel in a heterostructure-on-insulator (HOI) MOSFET in which a layer of strained $\text{Si}_{1-x}\text{Ge}_x$ is added in the middle which provides enhanced current compared to single s-Si on relaxed $\text{Si}_{1-x}\text{Ge}_x$ on insulator MOSFET. In [22], the drain current in a 50 nm channel length double s-Si with s-SiGe channel MOSFET increases by 41.3% in comparison with the 100 nm channel length MOSFET device.

The motivation of this paper is the study of the effect of a triple-layered strained Si channel consisting one $\text{Si}_{1-x}\text{Ge}_x$ layer sandwiched in between two silicon layers and having a mole fraction of 40% ($x = 0.4$) on a 3D trigate HOI n-channel FinFET and compare its electrical characteristics with another 3D TG FinFET employing only silicon as the channel.

2 Device Structure and Theory

Two TG FinFETs are designed involving two different channel materials. The details of the structures, which are characterized by the length (L_{Fin}), height (H_{Fin}), width (W_{Fin}) of the fin, and thickness of oxide layer, are listed in Table 1. The two structures are developed using identical dimensions and doping levels of the different materials, varying only in terms of materials used in the channel. While one of the devices is developed with only silicon as the channel material, the second device incorporates a trilayered channel consisting of a SiGe layer in between two layers of strained Si. Since silicon and germanium have a lattice mismatch of 4.2% [23], an alloy of SiGe with mole fraction of 40% of Germanium induced in the channel introduces strain in the layers at opposite sides of silicon layers in the channel. Thus, the device with three-layered channel is further characterized by the dimensions of the different materials in the channel. Thus, for the 14 nm HOI TG FinFET, both the s-Si layers in the channel are made 2 nm thick, and the SiGe thickness is fixed at 6 nm. Silvaco Atlas simulator is used in the modeling and characterization of the device (Fig. 1).

Both the SOI and HOI devices are modeled and characterized by using Silvaco TCAD tools [24]. The electrical parameters of the devices are also simulated using Silvaco Atlas simulator. In the simulation of the devices, the Shockley–Read–Hall (SRH) recombination using fixed lifetimes model, the band gap narrowing model, and Auger recombination models are also included for better modeling of the 14 nm gate length FinFETs.

While the gate length of both the structures is maintained at 14 nm, length, widths, and heights of the source and drain are fixed at 10 nm only thereby developing a square device. The thickness of the gate oxide layer is considered as 1 nm which makes both the SOI and HOI device identical in terms of geometric dimensions and doping concentrations so as to provide an equivalent comparison.

Table 1 Characteristics of the 14 nm TG FinFETs

Notations	Description	Dimension
L_D, L_S	Length of drain/source	10 nm
L_G	Length of channel	14 nm
$T_{\text{ox}} (\text{SiO}_2)$	Lateral oxide thickness	1 nm
W_{FIN}	Thickness of silicon fin	10 nm
H_{FIN}	Height of silicon fin	10 nm
$T_{\text{s-Si}}$	Thickness of strained silicon in channel	2 nm
T_{SiGe}	Thickness of $\text{Si}_{0.6}\text{Ge}_{0.4}$ in channel	6 nm
N_A	Doping of channel	10^{15} cm^{-3}
N_D	Doping of drain/source	10^{18} cm^{-3}

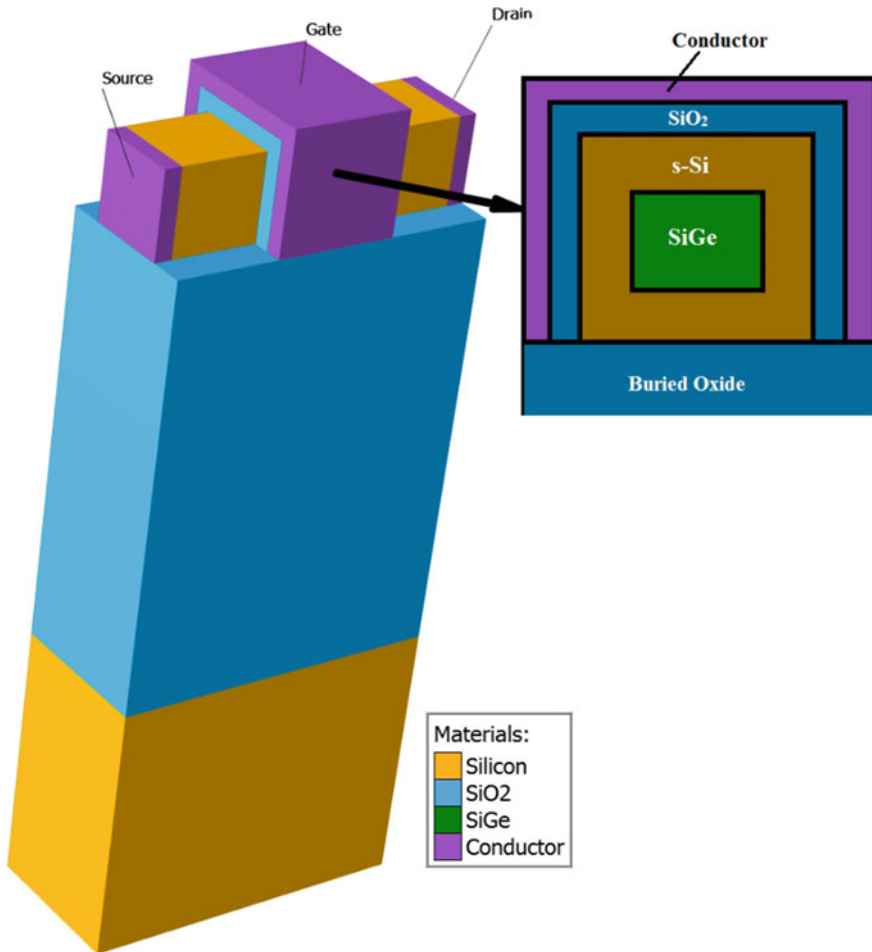


Fig. 1 Schematic of the HOI TG n-FinFET in 3D

3 Results and Discussion

The two FinFET structures are designed using Silvaco TCAD for gate oxide of 1 nm. The linear scale plot of $I_D - V_{GS}$ characteristics is shown in Fig. 2. The linear plot of the transfer characteristics determines the threshold voltages and the maximum on current, I_{on} , for the SOI and HOI device, respectively.

The variation of the threshold voltage is acquired for both the devices and presented in Table 2. From Fig. 2, the threshold voltage of the devices with silicon only channel and trilayered channel are determined to be 0.217 V and 0.256 V, respectively, with a drain voltage of 0.1 V. If the drain voltage is decreased to 0.05 V, then the threshold voltages increase to 0.226 and 0.265 V. Similarly, an increase in

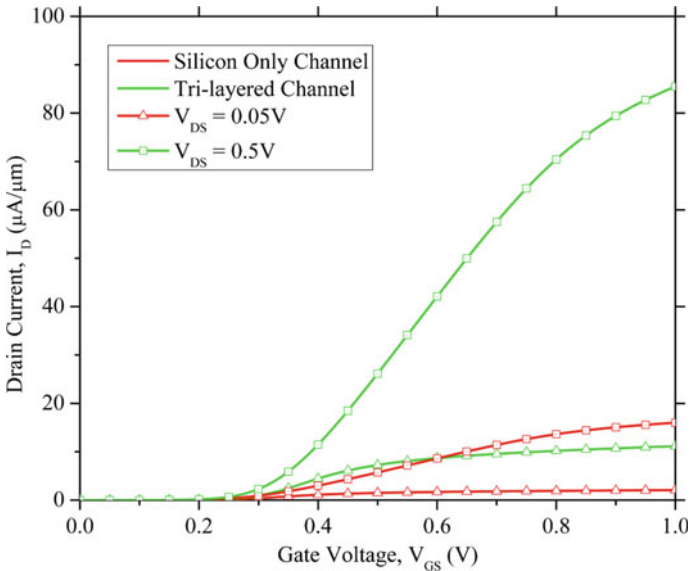


Fig. 2 $I_D - V_{GS}$ characteristics of TG FinFET with on a linear scale at $V_{DS} = 0.05$ and 0.5 V

Table 2 Variation of threshold voltages with change in V_{DS} voltage

Threshold voltage, V_{TH} (V)			
Drain voltage (V_{DS})	0.05 V	0.1 V	0.5 V
Silicon only channel	0.226	0.217	0.053
Trilayered channel	0.265	0.256	0.087

the drain voltage to 0.5 V shows a decrease in the threshold voltages to 0.053 and 0.085 V in both the devices, respectively, as shown in Fig. 3.

A large increase in drain current is observed in the HOI device with an increase in drain voltage as the gate voltage is varied from 0 to the supply voltage. For a drain voltage of 50 mV , the maximum drain current for the SOI device is $2.1\ \mu\text{A}/\mu\text{m}$, while it is $11.1\ \mu\text{A}/\mu\text{m}$ for the HOI device displaying an almost five times increase in drain current. A similar trend is also observed for increased drain voltages where the drain current for HOI device is $85.5\ \mu\text{A}/\mu\text{m}$ when compared to $16\ \mu\text{A}/\mu\text{m}$ for the SOI device at drain voltage of 0.5 V , and is given in Table 3. Hence, the HOI device holds an upper hand in respect to performance as shown in Fig. 4.

From Fig. 5, which shows the logarithmic plot of the $I_D - V_{GS}$ transfer characteristics, the leakage current, also known as the off current, the subthreshold swing, and the drain induced barrier lowering can be determined. The leakage current can also be determined by Eq. (1) as [25]:

$$I_{\text{off}}(nA) = 100 \frac{W}{L} 10^{-\frac{V_{TH}}{SS}} \tag{1}$$

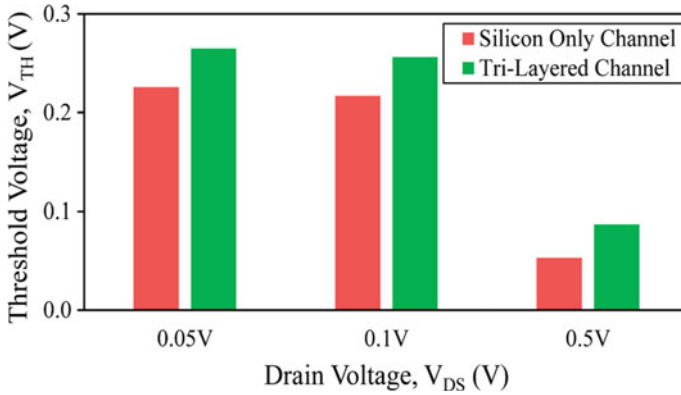


Fig. 3 Variation of threshold voltage (V_{th}) of a SOI and HOI TG n-FinFET

Table 3 Variation of I_{on} current with change in V_{DS} voltage

On current, I_{on} ($\mu A/\mu m$)			
Drain voltage (V_{DS})	0.05 V	0.1 V	0.5 V
Silicon only channel	2.1	4.02	16
Trilayered channel	11.1	21.71	85.5

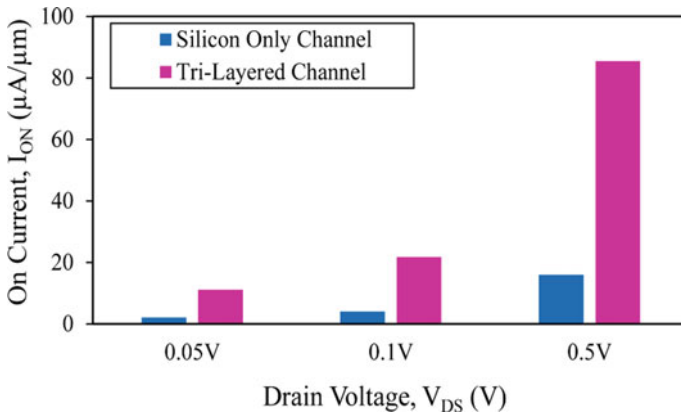


Fig. 4 Variation of on current (I_{on}) a SOI and HOI TG n-FinFET

where W refers to the width of the channel and L is the length of the channel. V_{TH} is the threshold voltage of the device, and SS is the subthreshold swing.

Table 4 shows that although the leakage currents increase with increasing drain voltages while sweeping V_{GS} from 0 to the supply voltage, the leakage currents are still in $nA/\mu m$ range also shown in Fig. 6, resulting in higher on/off current ratios. At a drain voltage of 50 mV, the FinFET with silicon only channel has a leakage

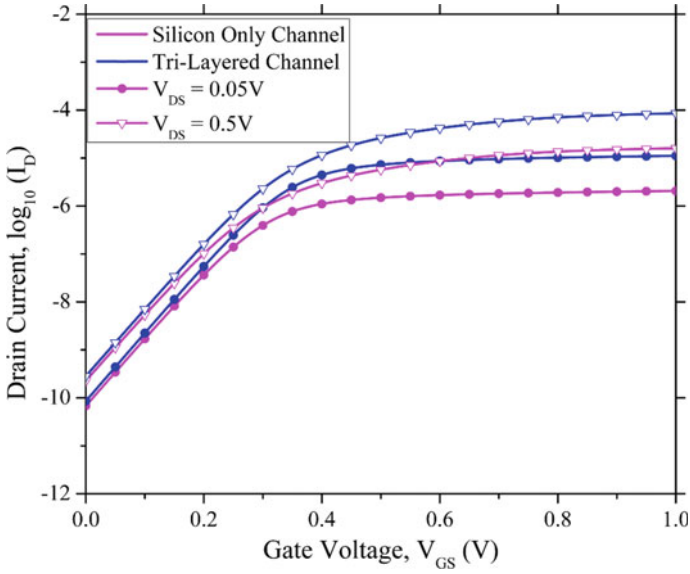


Fig. 5 $I_D - V_{GS}$ characteristics of TG FinFET on a log scale at $V_{DS} = 0.05$ and 0.5 V

Table 4 Variation of I_{off} current with change in V_{DS} voltage

Off current, I_{off} (nA/ μ m)			
Drain voltage (V_{DS})	0.05 V	0.1 V	0.5 V
Silicon only channel	0.068	0.089	0.228
Trilayered channel	0.086	0.112	0.284

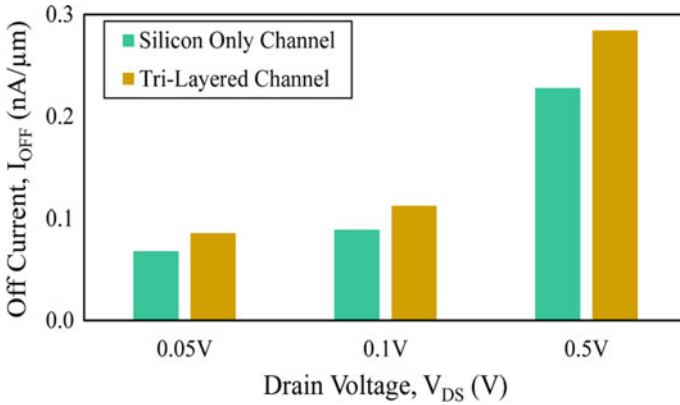


Fig. 6 Variation of off current (I_{off}) of a SOI and HOI TG n-FinFET

Table 5 Variation of I_{on}/I_{off} current ratio with change in V_{DS} voltage

I_{on}/I_{off} current ratio ($\times 10^4$)			
Drain voltage (V_{DS})	0.05 V	0.1 V	0.5 V
Silicon only channel	3.04	4.51	7.02
Trilayered channel	13.02	19.39	30.07

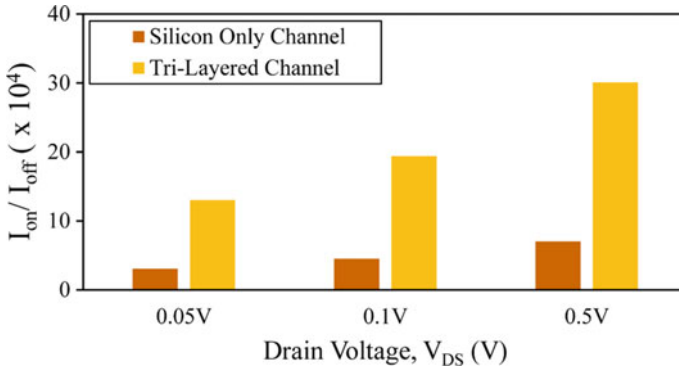


Fig. 7 Variation of I_{on}/I_{off} ratio of a SOI and HOI TG n-FinFET

current of $68 \text{ pA}/\mu\text{m}$, and the three-layer channeled device has a leakage current of $86 \text{ pA}/\mu\text{m}$ leading to an I_{on}/I_{off} ratio of 3.04×10^4 and 13.02×10^4 , respectively.

The subthreshold swing is observed to be around $71 \text{ mV}/\text{decade}$ for the SOI device with silicon only channel and around $70 \text{ mV}/\text{decade}$ for the strained silicon HOI device. A decade is a ten times change in the drain current.

From Table 5, it can be concluded that the I_{on}/I_{off} ratio of the strained silicon heterostructure device is more than four times the ratio of the silicon channel device. Figure 7 shows this comparison of I_{on}/I_{off} current ratio for both devices. This trend of the on/off current ratio is also observed for other drain voltages of 0.1 and 0.5 V thereby proving that the HOI structure has a higher switching speed and faster operation.

The DIBL, measured from the logarithmic plot of the $I_D - V_{GS}$ graph, is a short channel effect that calculates the amount of change in threshold voltage as the drain voltage is subjected to a large change. For improvement in the operation of the device, the DIBL should be low. The DIBL for both of the devices can be calculated using the standard Eq. (2) as [26]:

$$\text{DIBL} \left(\frac{mV}{V} \right) = \frac{\nabla V_{TH}}{\nabla V_{DS}} \tag{2}$$

where ∇V_{TH} refers to the change in the threshold voltage and ∇V_{DS} is the change in the drain to source voltage.

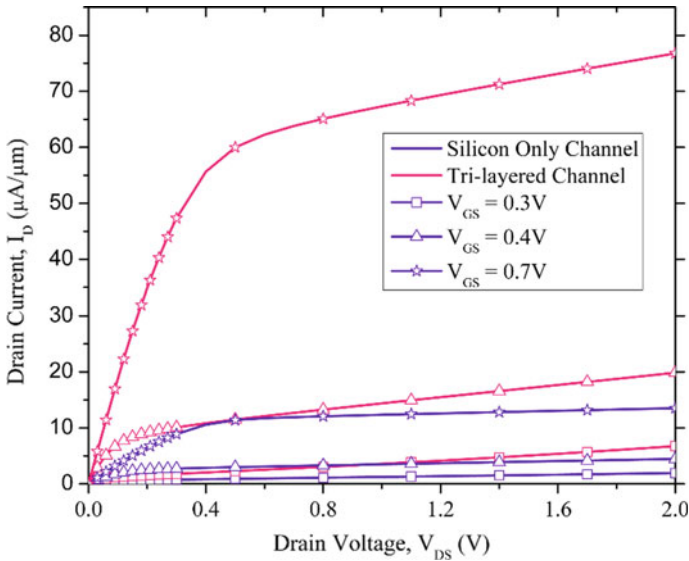


Fig. 8 $I_D - V_{DS}$ characteristics of a SOI and HOI TG n-FinFET at $V_{GS} = 0.3, 0.4$ and 0.7 V

The DIBL is determined as 83 mV/V for silicon only channel device and 58 mV/V for the device employing strained silicon channel, thereby confirming the better performance of the strained device.

Figure 8 shows the $I_D - V_{DS}$ characteristics of the two triple gate n-FinFET devices with different channel regions, for different gate-to-source voltages of 0.3 , 0.4 , and 0.7 V. The graph shows a sharp linear region and the early acquisition of saturation region of both the devices. Although the saturation current of the SOI device varies negligibly, the saturation current of the HOI device shows a small positive slope depicting better performance.

4 Conclusion

This paper describes the development of two trigate n-channelled FinFETs with 14 nm gate length using Silvaco TCAD software. While one device has only silicon as the channel material, the other device employs a three-layered s-Si/SiGe/s-Si as channel. The devices are developed with a height and width of 10 nm for implementation of a square device having small device area. The gate oxide layer of the devices are kept at 1 nm of SiO_2 . The threshold voltage, on current, leakage current, subthreshold swing, and DIBL are compared for both the devices for different gate and drain voltages. The I_{on} current for the silicon only channel device is 4.02 $\mu\text{A}/\mu\text{m}$ and for the strained silicon channeled device is 21.71 $\mu\text{A}/\mu\text{m}$ showing a five times increase, and the I_{on}/I_{off} ratio also shows a four times increase for the device with strained silicon in

the channel for very small increase in leakage currents. Thus, the TG FinFETs show improved performance with strained silicon channel and can possibly be downscaled below 14 nm technology node.

Acknowledgements The authors are grateful for the contribution of all the members of the Department of Electronics and Communication Engineering, NIT Mizoram, for providing the essential facilities like workstations and Silvaco TCAD tools as well as guidance throughout this work.

References

1. A. Bakibillah, N. Rahman, Effect of threshold voltage and channel length on drain current of silicon n-MOSFET. *ESJ* **11**(27) (2015)
2. J.K. Saha, N. Chakma, M. Hasan, Impact of channel length, gate insulator thickness, gate insulator material, and temperature on the performance of nanoscale FETs. *J. Comput. Electron.* **17**, 1521–1527 (2018)
3. A. Chaudhry, M.J. Kumar, Controlling short-channel effects in deep-submicron SOI MOSFETs for improved reliability: a review. *IEEE Trans. Device Mater. Reliab.* **4**(1), 99–109 (2004)
4. Y. Husaini, M.H. Ismail, A.S. Zoofakar, N. Khairudin, Electrical characteristics comparison between partially-depleted SOI and n-MOS devices investigation using Silvaco, in *2010 IEEE Symposium on Industrial Electronics and Applications (ISIEA)*, Penang, 2010, p. 532–536
5. J.P. Colinge, *FinFETs and Other Multi Gate Transistors* (Springer, New York, USA, 2008)
6. D. Hisamoto et al., FinFET—a self-aligned double-gate MOSFET scalable to 20 nm. *IEEE Trans. Electron Devices* **47**(12), 2320–2325 (2000)
7. N. Paydavosi et al., BSIM-SPICE models enable FinFET and UTB IC designs. *IEEE Access* **1**, 201–215 (2013)
8. A. Chaudhry, J.N. Roy, G. Joshi, Nanoscale strained-Si MOSFET physics and modeling approaches: a review. *J. Semicond.* **31**, 104001 (2010)
9. K. Rim et al., Fabrication and mobility characteristics of ultrathin strained Si directly on insulator (SSDOI) MOSFETs. *IEDM Tech. Dig.* **49**, 49–52 (2003). <https://doi.org/10.1109/IEDM.2003.1269163>
10. A. Wei et al., Integration challenges for advanced process strained CMOS on biaxially-strained SOI (SSOI) substrates. *ECS Trans.* **6**(1), 15–22 (2007)
11. C. Auth, A. Cappellani, J.S. Chun et al., 45 nm high-k+ metal gate strain-enhanced transistors, in *Symposium on VLSI Technology Digest, 2008*, pp. 128–129, ISSN 2158-9682
12. K. Ota et al., Novel locally strained channel technique for high performance 55 nm CMOS, in *IEDM Technical Digest, 2002*, pp. 27–30
13. C.H. Chen et al., Stress memorization technique (SMT) by selectively strained-nitride capping for sub-65 nm high performance strained Si device application, in *Symposium on VLSI Technology Digest, 2004*, pp. 56–57
14. P. Banerjee, P. Saha, D.K. Dash, A. Ghosh, S.K. Sarkar, Analytical modeling and performance analysis of graded channel strained dual-material double gate MOSFET, in *4th International Conference on Computing Communication and Automation 2018 (ICCCA)*
15. H.S. Yang et al., Dual stress liner for high performance sub-45 nm gate length SOI CMOS manufacturing, in *IEDM Technology Digest, 2004*, pp. 1075–1077
16. S. Gannavaram, N. Pesovic, M.C. Ozturk, Low temperature (<800 °C) recessed junction selective silicon-germanium source/drain technology for sub-70 nm CMOS, in *IEDM Technology Digest, 2000*, pp. 437–440
17. E.A. Fitzgerald, M.L. Lee, C.W. Leitz, D.A. Antobiadis, MOSFET Channel Engineering Using Strained Si, SiGe and Ge Channels (MIT, Cambridge, 2003). [Online]. Available: <http://hdl.handle.net/1721.1/3726>

18. P.K. Tiwari, G.K. Saramakala, S. Dubey, A.K. Mukhopadhyay, Analytical model for subthreshold current and subthreshold swing of short channel double material gate MOSFETs with strained-silicon channel on silicon-germanium substrates. *J. Semicond.* **35**(10) (2014)
19. S.H. Olsen, K.S.K. Kwa, L.S. Driscoll, S. Chattopadhyay, A.G. O'Neil, Design, fabrication and characterization of strained Si/SiGe MOS transistors. *IEE Proc. Circuits Devices Syst.*, **151**(5) (2004)
20. Sarah H. Olsen et al., Study of single- and dual-channel designs for high performance strained-Si-SiGe n-MOSFETs. *IEEE Trans. Electron Devices* **51**(7), 1245–1253 (2004)
21. L. Khiangte, R.S. Dhar, Development of double strained Si channel for heterostructure on insulator MOSFET, in *2017 2nd International Conference on Man and Machine Interfacing (MAMI)*, Bhubaneswar, 2017, pp. 1–3
22. L. Khiangte, R.S. Dhar, Double strained Si channel heterostructure on insulator MOSFET in sub-100 nm regime, in *2017 2nd International Conference on Man and Machine Interfacing (MAMI)*, Bhubaneswar, 2017, pp. 1–4
23. M. Jagadesh Kumar, T.V. Singh, Quantum Confinement Effect in Strained Silicon MOSFET. *Int. J. Nanosci.* **7**(2), 81–84 (2008)
24. S. International, *Atlas User's Manual Device Simulation Software* (Silvaco Int., Santa Clara, 2016)
25. S.K. Mohapatra, K.P. Pradhan, P.K. Sahu, Some device design considerations to enhance the performance of DG-MOSFETs. *Trans. Electr. Electron. Mater.* **14**(6), 291–294 (2013)
26. H.E. Ghitani, DIBL coefficient in short-channel NMOS transistors, in *Proceedings of the Sixteenth National Radio Science Conference. NRSC'99* (IEEE Cat. No. 99EX249), Cairo, Egypt, pp. D4/1–D4/5, 1999

Economic and Environmental Analysis of a Solar-Powered EV Charging System in India—A Case Study



Prashant Shrivastava , TeyKok Soon , Saad Mekhilef ,
and Furkan Ahmad 

Abstract With the growing environmental concern, the demand for electric vehicle is increasing in India. However, the limited existing charging facilities slow down the rate of adoption in the market. Presently, the solar-based charging system is gaining popularity, due to its low greenhouse gas emission. In this study, a case study for different solar charging systems suited in Delhi is presented. The economic and environmental analysis of different solar-based EV charging systems utilized for charging three types of commercial EVs such as 2, 3, and 4-W is explained. The cost of energy required for 100 km drive is lower with vehicle integrating PV charging system; however, it can act as a supportive system for driving range improvement. The analysis shows that the economically solar rooftop charging system is suitable for all types of EVs. The CO₂ emission produced by all types of solar-based EV charging systems is lower than coal-based charging systems.

Keywords Electric vehicle · Solar energy · Charging

P. Shrivastava (✉) · T. Soon

Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
e-mail: prashant@um.edu.my

S. Mekhilef

Power Electronics and Renewable Energy Laboratory, Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia

School of Software and Engineering, Faculty of Science, Engineering and Technology, Swinburne University of Technology, Melbourne, VIC 3122, Australia

S. Mekhilef

e-mail: saad@um.edu.my

F. Ahmad

Centre for Automotive Research and Tribology (CART), Indian Institute of Technology, New Delhi, Delhi 110016, India

1 Introduction

In India, emissions from thermal power plants are significant. Among all the countries in the world, India comes in third place after America and China in the biggest greenhouse gases (GHGs) emitters list [1]. According to the National Thermal Power Corporation (NTPC), about 58% of the installed power generation capacity of 331,117.58 megawatts (MW) is generated by coal [2]. As reported by the Center for Science and Environment (CSE) in 2017–18 [3], a significant amount of emission comes from the power sectors in industries in India. In the last decade, to counter the emissions from the thermal power plant, the renewable energy capacity has increased from 10 to 59 GW, and the government has set a target to achieve 100 GW by 2022 [4, 5].

In Indian metro cities, nearly 52% comes from industrial point sources such as power plants and 36% from vehicular emissions. So that the transportation sector is the second leading source of emission in India. Central and state governments in India have taken a series of steps to address the issue to control air pollution. Due to the change in government policies and new tax reforms for electric vehicle manufacturers, electric vehicles (EVs) are rising as succeeding options to conventional internal combustion engines [6].

The central government has allotted significant funding and taken several legislative and regulatory steps to promote EV deployment under the faster adoption and manufacturing of hybrid and electric vehicles mission (FAME-India mission) [7]. The growth and public adoption of EVs and PEVs are highly dependent upon the Li-ion battery development and on the development of charging facilities for these vehicles, to meet their power requirements. According to a survey conducted in India [8], the deployment of a comprehensive charging foundation with advanced equipment is a primary concern for EV users. However, the charging of EV from electricity generated from the thermal power plant produces more emissions than conventional IC engine vehicles. Furthermore, the charging from conventional power grid power supply cannot be considered as an environment-friendly method of charging [9].

Due to the high potential of solar energy in India, the solar-based EV charging system can play an important [7, 9]. With the application of solar for EV charging, the increase in the load on the conventional power grid can be significantly reduced. The solar-based charging can be two types off-grid and on-grid [10]. Mainly the off-grid solar-based EV charging systems can be divided in three categories, for example, solar carport [8], solar rooftop [11, 12], and vehicle-integrated photovoltaic (VIPV) [13, 14]. In this work, a case study is presented for different types of solar-based EV charging systems suited in Delhi city. Three types of commercial EV models are considered. Finally, the economic and environmental analysis of different solar-based EV charging system utilized for charging of considered EV models is presented.

In this study, three different types of commercial EVs are considered. The solar realistic meteorological data of Delhi is taken from the National Aeronautics and Space Administration (NASA)-based weather forecasting system which is used. The economic and environmental analysis of three solar-based EV charging systems are

performed. The running cost of EVs charged from different solar-based charging system for 100 km driving range is compared.

The paper is organized as follows. In Sect. 2, solar potential and EV market penetration in India are discussed. Sect. 3 describes the different solar-based EV charging systems. Case study for economic and environmental analysis of considered city in India is presented. Section 5 gives the conclusions.

2 Solar Potential and EV Market Penetration in India

2.1 Solar Potential in India

Presently, solar energy is playing a prominent role in the Indian electricity sector. Due to the high solar receiving capability of 4–7 kWh per sq. m per day in India, a great amount of solar energy can be produced, for example, 5000 trillion kWh per year [15]. Solar photovoltaics power can effectively be harnessed providing huge scalability in India. Solar also provides the ability to generate power on a distributed basis and enables rapid capacity addition with short lead times. The Indian government has considerably enlarged its solar plants with a target of installed capacity of solar to 100 GW by 2022 [5]. The state-wise targeted solar capacity set by the Indian Govt. to reach the 100 GW solar capacity is shown in Fig. 1. Due to large areas and maximum solar radiation throughout the year, Rajasthan and Uttar Pradesh have maximum potential to harness more solar energy. The targeted solar capacity for Rajasthan and Uttar Pradesh is 10,697 MW.

2.2 EV Market Penetration

In contrast to international EVs markets such as the USA, China, and Europe, the Indian EVs market is in its initial stage. It is less than 1% of the total EV sales; however, it the potential to grow to more than 5% in a few years. To pump up this, Indian Govt. is working on different missions such as the FAME scheme, in which, the central government has set the target to achieve production of around 7 Million EVs by 2020. Recently Indian Govt. has approved the FAME II mission with an outlay of Rs. 10,000 Crore for three years starting from April 1, 2019. The total number of EVs sold in different states of India by November 24, 2020, is presented in Fig. 2.

At present, there are more than 0.4 million 2-W, around 50 million of 3-W, and a few thousand 4-W on Indian roads. However, the industrial volumes change according to the incentives provided by the Govt. Currently, around 95% of EVs are low speed, for instance, electric scooters/rickshaws (less than 25 km/h) that do not require registration and licenses. Almost all the existing EVs run on lead batteries,



Fig. 1 Indian Govt. targeted solar power capacity in India by 2022 [5]

which keep the overall cost low, though low cycle life and frequent battery failure are the major roadblocks for the sales of EVs along with government subsidies. The projected EVs sale in India by 2030 is shown in Fig. 3. The sale of 2-W, 3-W, and 4-W will be reached to 5–7 million, 0.4–0.5 million, and 1–2 million, respectively.

2.3 EV Charging Tariff in India

The flat tariff has been announced by the Indian states and union territories (UTs) regularity commission which is presented in Fig. 4. The EV charging tariffs in India varies in the range of ₹ 4/kWh to ₹ 6/kWh. In some of the states, the charging

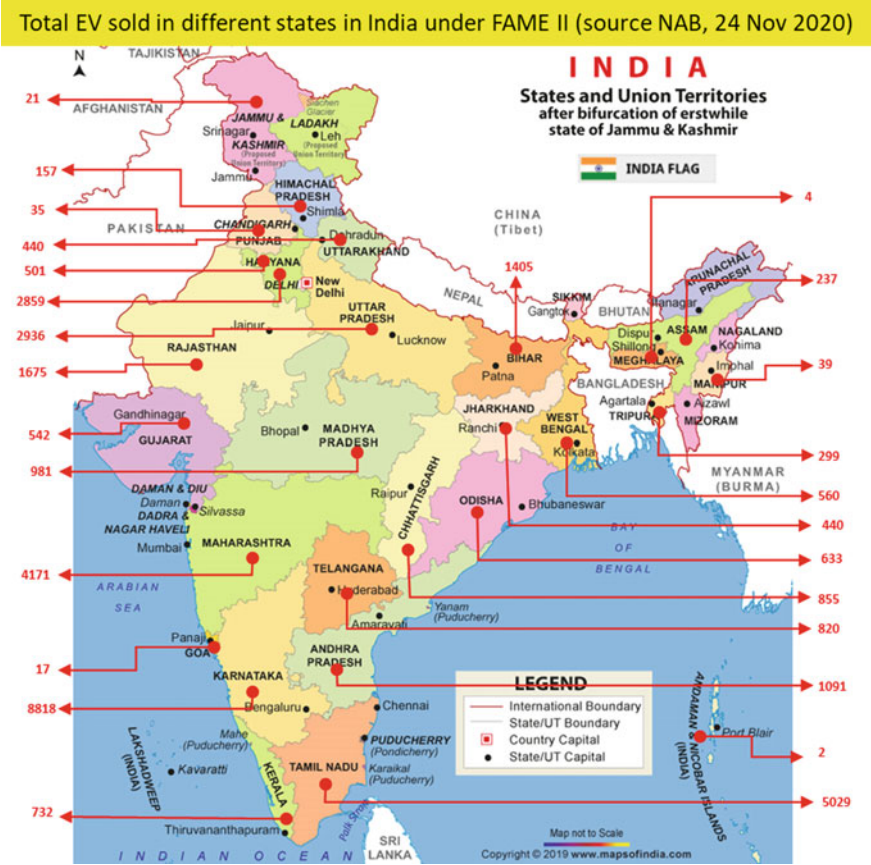
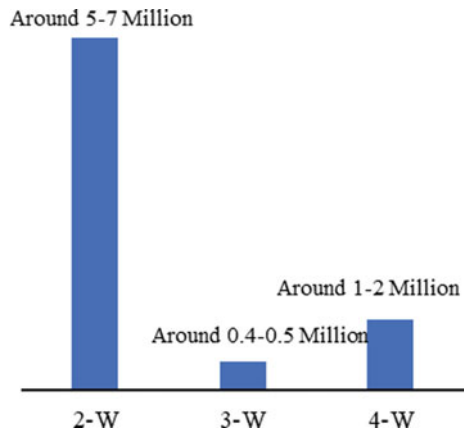


Fig. 2 Total EV sold in India under FAME II mission by Nov 24, 2020 [15]

Fig. 3 Projected EV sales in India by 2030 [16]



3 Different Solar-Based EV Charging Systems

Presently, three types of systems are greatly utilized to develop the solar-based EV charging system, for example, solar carport, solar rooftop, and vehicle-integrated PV (VIPV).

3.1 Solar Carport

A solar carport is a ground-mount structure type solar-based EV charging system. It also can be accommodated as parking for 2, 3, and 4-W. It is the most widely used solar-based charging system. It is suitable to decrease the heap and discharges from the power network yet additionally influences the expense of charging. This sort of charging system is reasonable for working environment parking structure where vehicles are left during the daytime. Particularly in India, representatives of most of the littler Indian organizations and MNCs are by and large labor for 11 h per day and 6 days every week. These perceptions show that vehicles are ordinarily left for an adequately lengthy timespan during the day to such an extent that it is conceivable to utilize solar-based vitality straightforwardly to charge the EVs in the work environment without utilizing enormous vitality stockpiling. As shown in Fig. 5, the expense of development to construct a parking space is a lot higher than a housetop establishment; notwithstanding, they have the preferred position that the transmission misfortunes become lower and give shade to the vehicle. The largest solar carport is installed by Tata in Kochi of 2.67 MW [18].



Fig. 5 Representation of solar carport

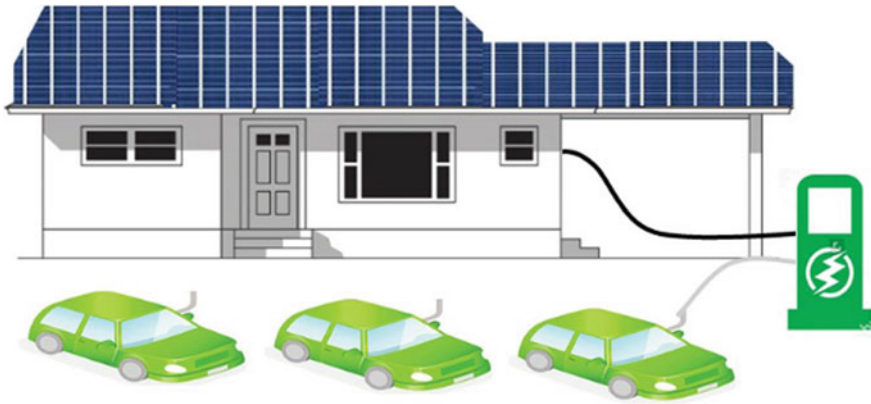


Fig. 6 Schematic of solar rooftop \times EV charging

3.2 Solar Rooftop

As shown in Fig. 6, the solar rooftop system comprises the roof of residential or commercial buildings covered by solar panels. The solar rooftop system comprises of different components such as solar inverters, PV modules, solar inverter, mounting system, and other electrical accessories. Compared to a solar carport which is in megawatt capacities, the rooftop system can be installed at low capacity. For example, the solar PV system on residential buildings typically ranges from 5 to 20 kW. Other side, the capacity of solar PV mounted on the commercial buildings often reaches 100 kW or more. The most cost-efficient solution is to place the panels on a roof of a nearby building which eliminates a shadow from trees, buildings, and other objects. The private companies for working together to develop the solar rooftop EV charging system, for example, Tata Power and MGL collaborate to develop solar rooftop EV charging system [19]. First Indian solar rooftop charging system installed by BHEL at Chandigarh highway in Delhi [20].

3.3 Vehicle Integrated Solar PV

As shown in Fig. 7, VIPV is a strategy in which the PV board is mounted on the EV body. VIPV framework bears an invite extra wellspring of electric vitality not exclusively to charge the battery during vehicle stop yet during vehicle activity also [21]. It can give the extra capacity to electronic control units such as actuators, warming, ventilation, and cooling units and other electronic segments. Also, when the vehicle is left, the overabundance of electrical vitality can be utilized to supply backup flows just as ventilation frameworks to chill off the inside in hot atmospheres. The last would already be able to be found an extra gear in arrangement vehicles. A

Fig. 7 Schematic of VIPV

cooler inside can increment not just the client comfort (cooler inside when entering the vehicle) and diminishes fuel utilization and GHG outflows (less cooling work for the cooling) yet additionally secures the inside hardware (a less warm worry for trim and upholstery). For example, in [14], the VIPV system was utilized for e-rickshaw to improve the driving range. To improve the emission balance by the electrical commercial vehicles, the VIPV scheme will be utilized [22].

4 Case Study

4.1 Solar Potential in Delhi

In this study, the realistic meteorological data of Delhi is taken from the National Aeronautics and Space Administration (NASA)-based weather forecasting system. The latitude and longitude of Delhi city is 28.7041° N, 77.1025° E. Monthly solar radiation, horizontal surface, of Delhi city ranges from 3.49 to 6.91 kWh/m²/day. The annual average solar irradiance is 5.34 kWh/m²/day as shown in Fig. 8.

4.2 Specification of Considered Solar PV Module

The monocrystalline solar PV module of 300 W is considered in this study. The detailed specification of the considered PV module is listed in Table 1.

4.3 Specification of Considered EV Models

To conduct the case study, the three different types of EVs are considered of different OEMs. The detailed specification of the EVs is listed in Table 2.

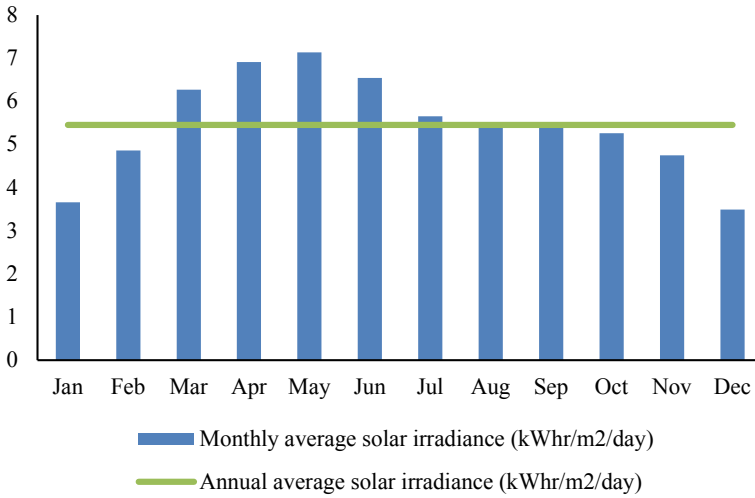


Fig. 8 Average solar irradiance of Delhi city horizontal surface by NASA

Table 1 Specifications of 300 W_p solar photovoltaic module [23]

S. No.	Parameter	Rating
1	Rated peak power	300 W
2	Fill factor	76.13%
3	Module efficiency	15.63%
4	Rated voltage (V_{mpp})	37.28 V
5	Rated current (I_{mpp})	8.05 A
6	Open circuit voltage (V_{oc})	44.1 V
7	Short-circuit current (I_{sc})	8.60 A
8	NOCT	45 °C
9	Dimension	1955 × 982 × 36 mm

4.4 Economic and Environmental Analysis

In this section, the economic and environmental analysis is performed. Three different types of solar-based charging systems are considered. The running cost of three types of EVs such as 2, 3, and 4-W charged from different solar-based charging systems is evaluated. Furthermore, the CO₂ emissions from different types EVs charged from considered different solar-based charging system for 100 km drive is calculated.

Cost of different solar-based charging systems

To evaluate the cost of the solar-based EV charging systems, different costs, for example, cost of PV panel, converter cost, maintenance cost, and installation charge, are considered. In this work, the cost of the PV panel is set to 55 ₹/W. As listed in

Table 2 Detailed specifications of considered EV models

Specifications	2-W (Hero Electric—Photon) [24]	3-W (Mahindra Treo Yaari -SFT) [25]	4-W (Mahindra e2o + P6) [26]
Motor power/type	1500 W/BLDC	850 W/BLDC	19 kW/3-phase induction motor
Battery rating/type	72 V, 24 Ah/VRLA	48 V, 100 Ah/VRLA	48 V, 210 Ah/lithium ion
Vehicle weight (kg)	114	215	1256
Maximum speed (kmph)	45	25	80
Charging time	4 h 30 m	2 h 30 m	6 h
Full-charge driving range (km)	110	130	110

Table 3, the cost of a solar carport, rooftop, VIPV is 1,23,000 ₹, 93,000 ₹, and 16,750 ₹, respectively. The cost of the VIPV charging method is the lowest; however, it cannot be implemented with 2 W. The cost of the solar carport is higher than rooftop due to its installation charge that includes the infrastructure cost also. The main benefit of the carport is that it can be utilized for parking.

Running cost of EV charged from solar-based charging system for 100 km drive

In Delhi city, based on the annual average solar irradiance per day of 5.34 kWh/m²/day, the energy production from 1 kW solar plant is 5.34 kWh/day. The life cycle of the solar module is 25 years. For 100 km drive, the calculated energy required using (1) by 2-W, 3-W, and 4-W is 1.73 kWh, 3.69 kWh, and 9.16 kWh, respectively.

$$\begin{aligned}
 &\text{Energy required per km (kWh)} \\
 &= \frac{\text{Battery pack energy capacity(kWh)}}{\text{Full charge driving range(km)}} \tag{1}
 \end{aligned}$$

Table 3 Cost of different types of solar-based EV charging systems

Specifications	Solar carport	Solar rooftop	VIPV
Power capacity	1 kW _p	1 kW _p	300 W _p
Cost of PV panel	55,000	55,000	13,750
Converter cost	18,000	18,000	1500
Maintenance cost	15,000	5000	500
Total installation charge (₹)	35,000	15,000	1000
Cost of charging station (₹)	1,23,000	93,000	16,750

The cost of energy production from different solar charging system can be calculate as below:

$$\text{Cost of energy production (\text{₹}/\text{kWh})} = \frac{\text{Cost of charging system (\text{₹})}}{\text{Per day energy production of solar charging system (kWh)} \times \text{life cycle of solar panel (days)}} \tag{2}$$

The total energy production of carport, rooftop, and VIPV in 25 years is 48,727, 48,727, and 14,618.25 kWh. As presented in Fig. 9, the calculated cost of energy production of a solar carport, rooftop, and VIPV are 2.52, 1.90, and 1.14 ₹/kWh. The calculated running cost of 2-W for 100 km drive using carport and rooftop charging system is 4.35 ₹ and 3.23 ₹. The calculated running cost of 3-W for 100 km drive using carport, rooftop, and VIPV charging system are 9.30 ₹, 7.01 ₹, and 4.20 ₹. The calculated running cost of 4-W for 100 km drive using carport, rooftop, and VIPV charging system is 23.10 ₹, 17.40 ₹, and 10.44 ₹. The running cost of EV with VIPV is the lowest; however, this type of charging system can be used for improving the driving range only. It cannot be used for the full charging of EVs.

CO₂ emission from EV charged from solar-based charging systems for 100 km drive

The CO₂ emission by EVs charged from a solar-based charging system is compared with the conventional vehicles for 100 km drive. According to [21], the CO₂ emission from the electricity generated by the solar system is 0.5 kg/kWh. As presented in Fig. 10, the calculated CO₂ emission for 100 km drive using (3) for considered EVs such as 2-W, 3-W, and 4-W is 0.865 kg, 1.845 kg, and 4.58 kg, respectively. As per the [22], the average CO₂ emission from thermal power plant ranges from 0.91 to 0.95 kg/kWh. Therefore, the calculated CO₂ emission using (3) from considered EVs charged from electricity generated by thermal power plants are 1.64 kg, 3.51 kg, and

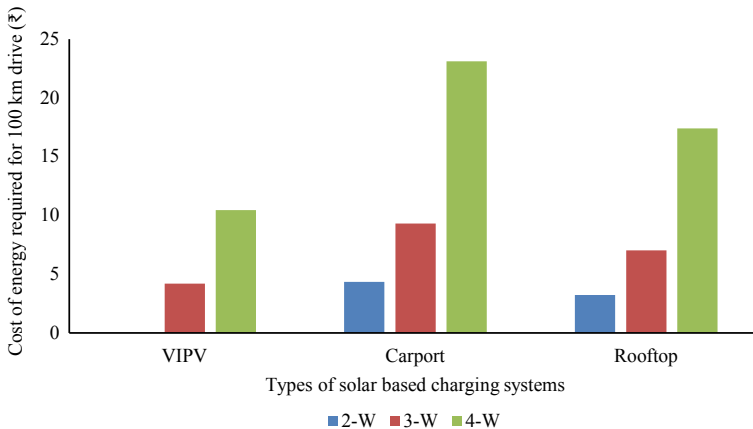


Fig. 9 Cost of energy required by considered EVs charged by different solar-based charging systems for 100 km drive

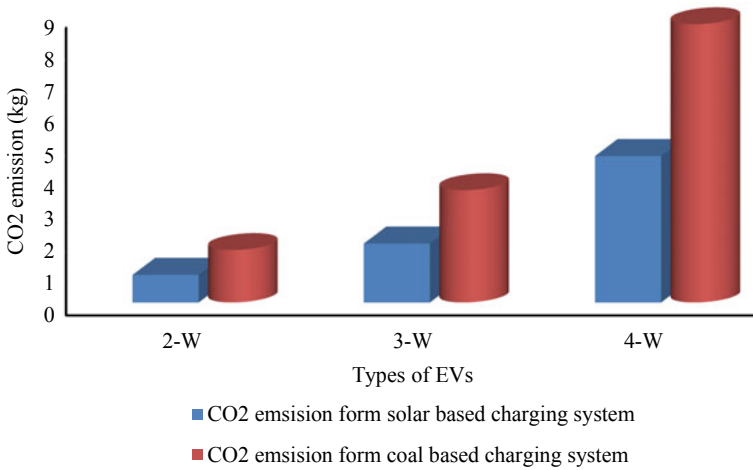


Fig. 10 CO₂ emission from considered EVs charged from electricity generated from solar and thermal (coal) power plant

8.702 kg for 2-W, 3-W, and 4-W, respectively. For all type of solar-based charging systems, the CO₂ emission is lower than thermal power (coal)-based charging system.

$$\begin{aligned}
 & \text{CO}_2 \text{ emission for 100 km drive(kg)} \\
 & = \text{Energy required for 100 km drive(kWh)} \\
 & \times \text{CO}_2 \text{ emission for unit generation of electricity(kg)} \qquad (3)
 \end{aligned}$$

5 Conclusion

For the efficient deployment of EVs in the market, the development of adequate charging infrastructure is always a crucial task. Due to the high potential of solar energy in India, solar energy can play an important role in the development of economic and environmentally friendly charging system. In this work, the three type of solar-based charging systems have been studied, for example, solar carport, rooftop, and VIPV. To analyze the economic and environmental impact of the solar-based charging system, the case study results have been presented. The results demonstrate that the VIPV is the most economic solar charging system; however, it cannot be utilized as the main source of charging of EV and not suitable for 2-W. Though, the solar rooftop charging system is economically suitable for all types of considered EVs. The CO₂ emission produced by solar-powered EVs is nearly half of the emission produced by the thermal plant (coal)-based charging systems.

Acknowledgements The authors would like to thank Ministry of Higher Education, Malaysia under Large Research Grant Scheme (LRGS): LRGS/1/2019/UKM/01/6/3 and University of Malaya, Malaysia for providing financial support under the research grant RU-SATU Joint Research 2020: ST009-2020.

References

1. 'India third largest contributor to carbon emission'—The Hindu. <https://www.thehindu.com/news/national/co2-levels-poised-for-record-high/article25674341.ece>. Accessed Nov 25 2020
2. G. Chaudhary, P. Shrivastava, M.S. Alam, Y. Rafat, Estimates of Emissions From Thermal Plants in India and Method To Reduce Down These Emissions and Running Cost of Air Conditioner, 2016, pp. 1–6
3. "Centre for Science and Environment." <https://www.cseindia.org/page/annual-reports>. Accessed Nov 25 2020
4. J. Charles Rajesh Kumar, M.A. Majid, Renewable energy for sustainable development in India: Current status, future prospects, challenges, employment, and investment opportunities. *Energy. Sustain. Soc.* 10(1), 1–36 (2020). <https://doi.org/10.1186/s13705-019-0232-1>
5. Solar Performance Tracker | Engaging Indian States. <https://indianstates.csis.org/national-goals/solar-performance-tracker/>. Accessed Nov 24 2020
6. P. Mohanty, Y. Kotak, Electric vehicles: status and roadmap for India. *Electr. Veh. Prospect. Challenges*, pp. 387–414 (2017). <https://doi.org/10.1016/b978-0-12-803021-9.00011-2>
7. FAME INDIA II Scheme: Department of Heavy Industry, Ministry of Heavy Industries & Public Enterprises, Government of India. <https://dhi.nic.in/UserView?mid=1378>. Accessed Nov 25 2020
8. M. Asaad, P. Shrivastava, M.S. Alam, Y. Rafat, R.K. Pillai, Viability of xEVs in India: a public opinion survey. *Lect. Notes Electr. Eng.* **487**, 165–178 (2018). https://doi.org/10.1007/978-981-10-8249-8_15
9. A.R. Bhatti, Z. Salam, M.J.B.A. Aziz, K.P. Yee, R.H. Ashique, Electric vehicles charging using photovoltaic: status and technological review. *Renew. Sustain. Energy Rev.* **54**, 34–47 (2016). <https://doi.org/10.1016/j.rser.2015.09.091>
10. S. Khan, A. Ahmad, F. Ahmad, M. Shafaati Shemami, M. Saad Alam, S. Khateeb, A comprehensive review on solar powered electric vehicle charging system. *Smart Sci.* **6**(1) 54–79 (2018). <https://doi.org/10.1080/23080477.2017.1419054>
11. S. Prajapati, E. Fernandez, Rooftop solar PV system for commercial office buildings for EV charging load (2019). <https://doi.org/10.1109/icsima47653.2019.9057323>
12. Charging Your Electric Car With A Solar Panel Rooftop | Cleanenergyauthority.com. <https://www.cleanenergyauthority.com/solar-energy-news/charging-your-electric-car-solar-panel-rooftop-060520>. Accessed Nov 25 2020
13. K. Chaudhari, A. Ukil, K.N. Kumar, U. Manandhar, S.K. Kollimalla, Hybrid optimization for economic deployment of ESS in PV integrated EV charging station. *IEEE Trans. Ind. Inf.* **3203**, 1, (2017). <https://doi.org/10.1109/tii.2017.2713481>
14. P. Shrivastava, M.S. Alam, M.S.J. Asghar, Design and techno-economic analysis of plug-in electric vehicle-integrated solar PV charging system for India. *IET Smart Grid* **2**, 224–232 (2019). <https://doi.org/10.1049/iet-stg.2018.0079>
15. Current Status | Ministry of New and Renewable Energy, Government of India. <https://mnre.gov.in/solar/current-status/>. Accessed Nov 24 2020
16. What Is The Future of Electric vehicles in India by 2030—Promoting Eco Friendly Travel. <https://electricvehicles.in/what-is-the-future-of-electric-vehicles-in-india-by-2030/>. Accessed Nov 24 2020
17. S. Das Bhawna Tyagi, Dissecting India Electricity Tariff Landscape for EV (2020)

18. India's Largest Solar Carport by Tata Power. <https://www.tatapower.com/plants-projects/solar-rooftop-projects/cochi-carport.aspx>. Accessed Nov 25 2020
19. Tata Power and MGL to Collaborate on Solar Rooftop and EV Charging Stations—Mercom India. <https://mercomindia.com/tata-mgl-collaborate-solar-ev-charging/>. Accessed Nov 25 2020
20. BHEL: BHEL installs first solar EV charging station on Delhi-Chandigarh highway, Energy News, ET EnergyWorld. <https://energy.economicstimes.indiatimes.com/news/power/bhel-installs-first-solar-ev-charging-station-on-delhi-chandigarh-highway/68251129>. Accessed Nov 25 2020
21. S. Mandal, S. Biswas, V.E. Balas, R. N. Shaw, A. Ghosh, Motion prediction for autonomous vehicles from Lyft dataset using deep learning, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, 2020, pp. 768–773. <https://doi.org/10.1109/iccca49541.2020.9250790>
22. Vehicle-integrated PV for electrical commercial vehicles | PV Europe. <https://www.pveurope.eu/solar-generator/vehicle-integrated-pv-electrical-commercial-vehicles>. Accessed Dec 19 2020
23. Tata Power Solar | TP300 Series 300-330 | Solar Panel Datasheet | ENF Panel Directory. <https://www.enfsolar.com/pv/panel-datasheet/crystalline/35321>. Accessed Nov 26 2020
24. Photon LP—Hero Electric. <https://heroelectric.in/photon/>. Accessed Nov 26 2020
25. Mahindra Treo Zor, Price & Features | Electric Three Wheeler Loader. <https://auto.mahindra.com/electric/treo-zor>. Accessed Nov 26 2020
26. Know about the Mahindra e2oPlus Car Engine & Mileage | Mahindra Electric. <https://www.mahindraelectric.com/vehicles/e2oPlus/>. Accessed Nov 26 2020

Islanding Detection Review Using Intelligence Classifier in Distribution Network



Hairul Nissah Zainudin, Saad Mekhilef, Hazlie Mokhlis, and Safdar Raza

Abstract Islanding detection method is the most crucial aspect in distribution network. Generally, it can be categorized as remote, passive, active, and hybrid methods. From all these methods, islanding preferable inclined to the passive method since it is cheaper and able to maintain a power quality of the system. There are few drawbacks of passive techniques such as a high non-detection zone and its required onset setting. In order to overcome the drawback and limitation, various signal processing and intelligent techniques are introduced. Intelligent classifier techniques are particular to resolve the issues including the detection accuracy, speed of detection, and compatibility of detecting islanding in hybrid distributed energy resources. This paper offers a general summary of conventional islanding detection methods and focuses on islanding detection using intelligence classifier methods. Intelligence classifier will increase capability of common passive islanding detection methods hence upgrading the signal processing techniques. A comparison between the intelligence classifier methods with an existing techniques is also provided. In conclusions, this paper summarizes advantages and disadvantages of the intelligence classifier techniques for providing initial strategies to those researchers or power engineers for them to select the best option for their system.

H. N. Zainudin (✉) · S. Mekhilef · H. Mokhlis

Power Electronics and Renewable Energy Research Laboratory (PEARL), Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia

S. Mekhilef

e-mail: saad@um.edu.my

H. Mokhlis

e-mail: [fazli@um.edu.my](mailto:hazli@um.edu.my)

S. Mekhilef

School of Software and Electrical Engineering, Faculty of Science, Engineering and Technology, Swinburne University of Technology, Melbourne, VIC 3122, Australia

S. Raza

Department of Electrical Engineering, NFC IET, Multan Punjab, Pakistan

Keywords Islanding detection methods · Distributed generators · Intelligence classifier techniques · Non-detection zone

1 Introduction

The demand of electricity has increased exponentially since few years. This demand is fulfilled by using conventional as well as renewable energy sources. The conventional sources include oil, natural gas, and coal. The higher price of fuel, effect of the greenhouse, and several of pollution issues become major concerns while using conventional sources [1, 2]. Kyoto protocol [3] and EU-28 (European Union-28) policy [4] has been signed by many countries, and this agreement targets to reduce the greenhouse gas release to atmosphere by the year 2050.

The usage of renewable energy sources replaces the consumption of fossil fuels. Thus, enhances the regional economic development by introducing new business models. The EU has agreed to reduce the greenhouse gas emission to 90–95 percentage by year 2050. The intermediate goals by the year 2020 and 2030 relative to 1990 are 20% and 40%, respectively [4]. The steady and dynamic behavior of renewable energy sources become an issue in management and system protection. Islanding is the main issue in this perspective [5]. An islanding is a state in which distribution system losses the grid supply but at the same time continues to be energized by associated renewable energy sources (RES) [6–8].

Islanding can be detected by many techniques which are generally considered into two types: remote and local techniques [9]. Remote islanding detection techniques are efficient as compared to local techniques. Remote techniques based on communication methods to which transmitting a data between the grid and the DGs. There is no power quality and network consistency effect in communication-based methods. The communication-based methods are power line carrier communication (PLCC), supervisory control, and data acquisition (SCADA) methods [10]. Unfortunately, this communication-based methods need an expensive equipments. Local islanding detection methods (IDMs) are characterized into passive, active, and hybrid methods. Passive methods monitor the local variables and compare them with predetermined onset values [10]. Passive method will not perturb the system but it cause a large non-detection zone (NDZ) [7]. In active methods, disturbances are injected to the system, and the response is monitored in order to decide whether islanding has occurred or not [11]. However, it suffers from power quality problem. Meanwhile, hybrid islanding method is a combination of both passive and active method. It has the advantages of both active and passive methods but still suffers from non-detection zone [12].

The issue of NDZ and threshold setting requirement in local techniques may get resolved by employing the intelligence classifier [13]. The intelligence techniques have the capability to remove the invisible structures in measurement. The measure signal without invisible structures is an advantage in IDMs. Aim of this paper is to analyze the IDMs using intelligent classifier in distribution network. This paper

been arranged initially with an overview of microgrid architecture and distributed energy can be substituted into it. Secondly, it is about an overview of IDMs conventional methods with advantages and drawback given. Followed by a popular signal processing techniques that embedded with conventional method. Lastly, discussion followed by intelligent classifier method together with signal processing tools to attain greater accuracy, quicker detection, and compatible in distribution network.

1.1 Microgrid (MG)

A technology, having a cluster of interconnected loads and distributed energy resource (DER) at a small scale, is called as MG. The key purpose of MG is to guarantee steady operation during faults and any other instabilities in the system [14, 15]. It is formed by integrating loads, DER, and batteries [16]. Batteries are used to stabilize the variance between power produced by DERs and that utilize by the consumer [17]. MG is relatively a new type of network structure capable of self-protection, control, and management [18].

MG continuously monitors the changes in the system and manages to supply power without interruption. MG main advantage is the best solution of power supply during an emergency and failure in the main grid. MG is able to function during connection or islanding mode existence in a grid [14].

Figure 1 shows an architecture of MG. Main objective of MG is to deliver an electricity to minor groups of consumer such as buildings, schools, industry, or hospital. MG used fossil fuels and renewable energy sources as its main source of electricity [19].

Energy sources can be classified into non-distributed and DERs as shown in Fig. 2. Non-DERs include nuclear energy, fossil fuel, and thermic (thermos) which also not renewable energy. Meanwhile, DERs are classified based on high-energy efficiency

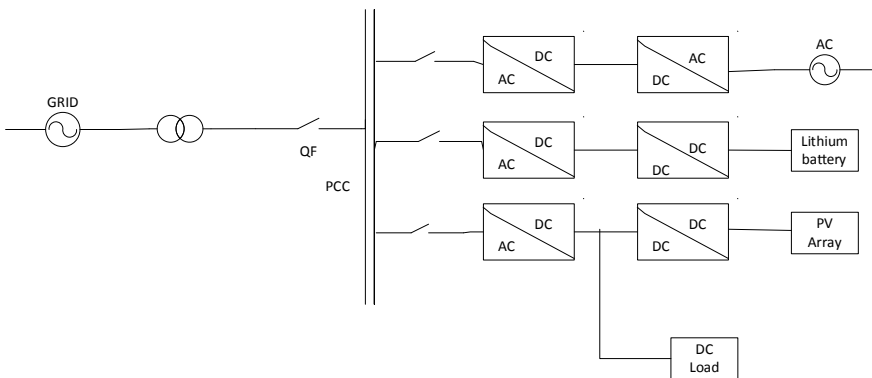


Fig. 1 Construction of MG

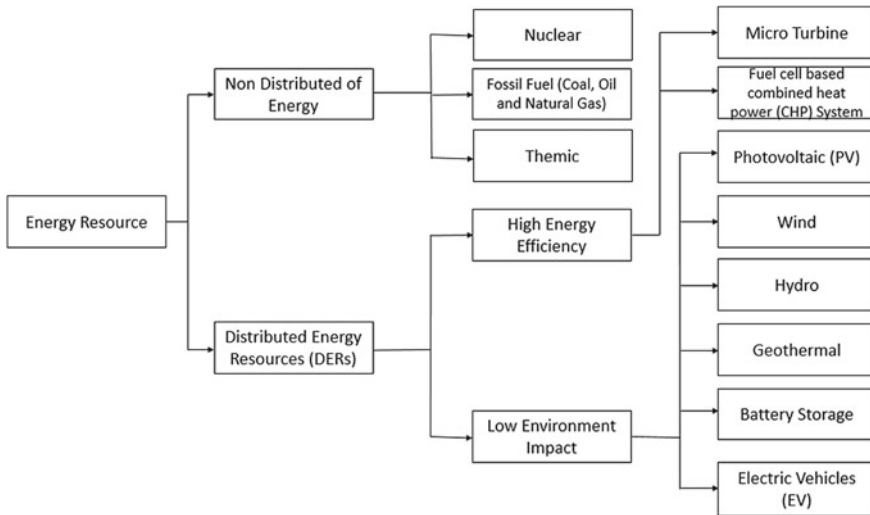


Fig. 2 Distributed and non-distributed energy resources

and minimal adverse impact on environment. Both types of DERs are able to produce high-energy efficiency but it exhausted a smoke to environment. However, DERs causing low environmental impact are photovoltaic (PV), wind, hydro, geothermal, and battery storage [11].

DERs improve the system reliability and decreases the need of conventional energy [19, 20]. Key feature of MG is when the synchronized operation and control can be done to DER, controllable loads, and storage devices. This is crucial element to make sure the MG may offer the maximum of its capacity [21].

1.2 Recent Trend of MG

MG is one of the best selection for hospitals, server, and the industries which requires uninterrupted power supply [22]. Efforts have been made by few countries such as Canada, USA, and Japan for MG research and implementation [21]. The research involved of few stages of in order to get the best research outcome. The first stage of research is exploration, improvement, and confirmation of process, mechanism, shield, protection, and telecommunication infrastructure of MGs. During this stage, the researchers investigate the validation of system operation without sacrificing the protection and safety of the system. Choosing an appropriate control and telecommunication infrastructure are most important during this stage [14].

In electrical power system, MG is facing few type of constraint such as penetration of DERs, characteristics of the load, and power quality. Unbalance of the active and reactive power will cause a deviation in system frequency [14].

The MG provides a control strategy between islanding and grid connection. Two types of control strategies of islanding are used to activate an inverter. During normal condition, each DER inverter in the MG system works in power or current control mode. DER will deliver a preset or available maximum power to the load. While in cut off of MG from the main grid condition, DER inverter needs to detect islanding. DER inverter will be switched to voltage control mode to provide a constant voltage to the load in order to maintain the stability of MG system [23].

Main challenge in MG implementation is its protection [24]. It is crucial to make sure the DERs, load, and lines DERs are protected [25]. Basically, MG consists of a cluster of feeder part in distribution network. There is a single point in this network which is known as point of common coupling (PCC). Those feeders will be feeding into a load which can be classified into constant impedance load, constant power load, and constant current load [23].

DERs are used in small-scale power system grid application. Generally, it increases the power reliability at usage load. DERs are able to deliver power if when utility grid failed to do so. Hence, DERs are able to eliminate transmission losses, reduce cost of transmission network, and reduce national grid system peak demand which results in lower initial capital [26, 27].

2 Islanding

Islanding operation in distribution system usually happens while the power supply from the main utility or grid is disconnected, however the DERs keep delivering power to the load [11, 28–31]. The power supply is disconnected with few reasons such as some unpredictable occasions: storms, earthquakes, tremendously hot weather, co-ordination failures, and human mistakes. The islanding is an uninvited occurrence since it causes a lot of effect to human life (electrical workers), the power quality in the system, equipment failure, and any other disturbance in the system [32]. Islanding can be classified into intentional or unintentional which is explained in Sect. 2.1 [33].

There are few basic requirements for a successful detection. Firstly, the scheme must detect any possible formations of islands. Secondly, the scheme should detect islanding conditions within the required time frame as specified in IEEE 1547 standard [34]. There are few international standards which exist besides IEEE 1547, and it is IEC 61727 and UL 1714 [35].

2.1 *Intentional and Unintentional Islanding*

Intentional islanding occurs during the maintenance service on the utility grid [7]. It often exists in industrial plants where the plant has excess electrical power for their smooth operation. For example, paper and sugar mills can produce electricity

Table 1 Islanding detection standard

Standard	Quality factor, Q_f	Islanding detection time, t	Frequency range	Voltage range
IEC62116	1	$t < 2$ s	$(f_o - 1.5 \text{ Hz}) < f$ and $(f_o = 1.5 \text{ Hz})$	$85\% < V < 115\%$
UL 1741	2.5	$t < 2$ s		$88\% < V < 110\%$
IEEE 1547	1	$t < 2$ s	$59.3 \text{ Hz} < f < 60.5 \text{ Hz}$	$88\% < V < 110\%$
Korean std	1	$t < 0.5$ s	$59.3 \text{ Hz} < f < 60.5 \text{ Hz}$	$88\% < V < 110\%$
IEEE929-2000	2.5	$t < 2$ s	$59.3 \text{ Hz} < f < 60.5 \text{ Hz}$	$88\% < V < 110\%$
VDE 0126-1-1	2	$t < 0.2$ s	$59.3 \text{ Hz} < f < 60.5 \text{ Hz}$ $47.5 \text{ Hz} < f < 50.5 \text{ Hz}$	$88\% < V < 110\%$

which fulfill their internal need. During thunderstorm or any unpredictable weather, this design is able to switch to internal production and segregate themselves from the main grid. By this phenomena action, it lowers the risk of instabilities because of lightning or any other faults [33, 36]. One more example of intentional islanding is an emergency backup in the hospital. To ensure a constant power delivery, most of hospitals having an emergency generators that are capable to provide power to each crucial equipment at any time in any condition [33].

Unintentional islanding happens when the grid is disconnected due to line tripping, equipment failure, and human error [10, 37]. This causes more deviations in voltage and frequency [33]. Generally, the islanding condition occurs at medium and low voltage distribution system.

There are few international standards in islanding for independent power producers (IPP) to have guideline before they can do a design and operate the island [29]. Table 1 summarizes the standard with their main parameters. Different country will apply different standard for their grid system. Tenaga National Malaysia (TNB) used IEEE 1547 standard for their MG system [26].

3 Islanding Detection Methods (IDMs)

IDMs are generally divided into four mean methods: local, remote, signal processing, and intelligent methods as shown in Fig. 3 [10, 38–43]. Local methods are usually related with DERs, and the remote methods are related with the utility side [44]. Local methods can be categorized into passive, active, and hybrid methods. In all these categorized, passive method is more valuable and offers minimum power quality degradation, less expensive, and commonly use in power utilities [43] (Fig. 4).

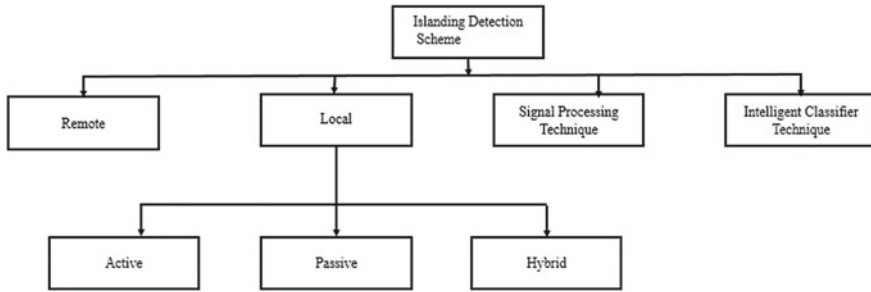


Fig. 3 General classification of IDMs

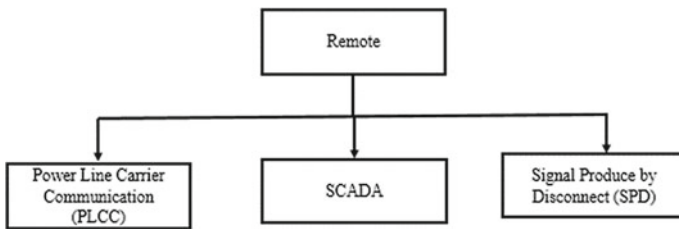


Fig. 4 General classification of remote methods [10]

3.1 Remote Methods

Remote techniques are the most effective and expensive method for islanding detection. In power line carrier communication (PLCC), it has transmitter and the receiver side. The transmitter is usually been located in a grid while the receiver will be in DER side. PLCC signal has been designed with four consecutive cycles, if the signal becomes weaken in three consecutive cycles which mean that islanding occurs. This means that there are interruptions in the signal. Usually, this detection will take around 200 ms. There will be no non-detection zone if the system is having a normal load [10]. Advantages of PLCC are that it can maintain the power quality of the system with zero impact on grid transient response. PLCC is highly recommended if the system is having in multiple-inverter. This method is not economical in low-density DER systems because transmitters are much expensive. Thus, this technique is advisable to apply in the MG having high-density DERs [10].

Another method is called as signal produce by disconnect (SPD). In SPD, islanding has been detected through the signal transmission between inverters of DER and external power grid. SPD signal transmission is based on microwave, telephone line, and other forms. SPD use the consecutive carrier signal, to avoid any failure caused by generator, network or receiver. The main advantage of SPD is it eliminates all NDZ. But SPD method need is high initial cost since we need to bear a cost of the communication wire and transmit signal and repeaters since we need telephone

line and microwave, respectively [10]. All this communication tools allocate extra charges to the whole islanding big system.

Another remote method is supervisory control and data acquisition (SCADA). It monitors the state of main grid [45]. In the case of islanding, it sends signal to corresponding DER. SCADA uses a same procedure as SPD but it is having extra control of DERs. NDZ can be eliminated from the system with higher efficiency if it has proper installation and good communication network. But during busy network, the detection for islanding is very slow in SCADA. Since we need individual device and communication link in multiple inverter case, this method considers expensive. Furthermore, complex installation of the system and certification, this technique is not appropriate for moderate system. Complex installation and certification showed that skill full workers need to be involved, and extra charges also need to be include [10].

3.2 Passive Methods

Passive IDMs are using local measurements of power system parameters such as frequency, voltage, phase angle, and total harmonic distortion detection at the point of common coupling (PCC). Passive methods detect islanding conditions by measuring changes in the electrical quantities at the DER output. Passive methods are cheap and simple due to reduced complexity and maintain the quality of power [31]. However, passive methods are less effective compared to active methods in detecting island operation due to their large NDZ [38, 46, 47].

Threshold setting requirement is also one of the limitations of passive islanding detection technique especially during closely power mismatch conditions. If the power mismatch is too close, less number of threshold can be set. This will affect the final result for islanding detection.

Figure 5 illustrates a principle of passive IDM. Passive techniques monitor the signal parameters such as voltage, frequency, harmonic distortion, and current on the DER side at the PCC [48]. If the parameter having higher threshold value, it means islanding occurs in the system and the DER should be disconnected immediately.

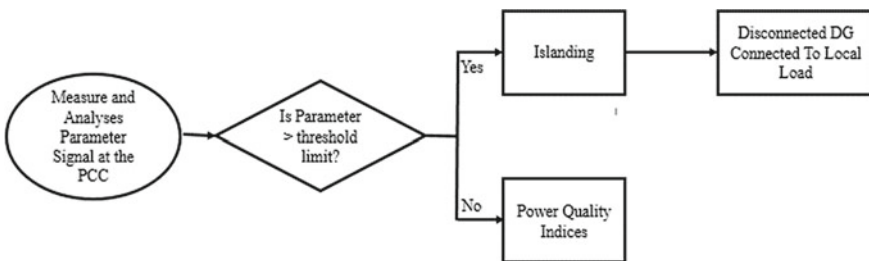


Fig. 5 Basic working principle of passive islanding detection technique

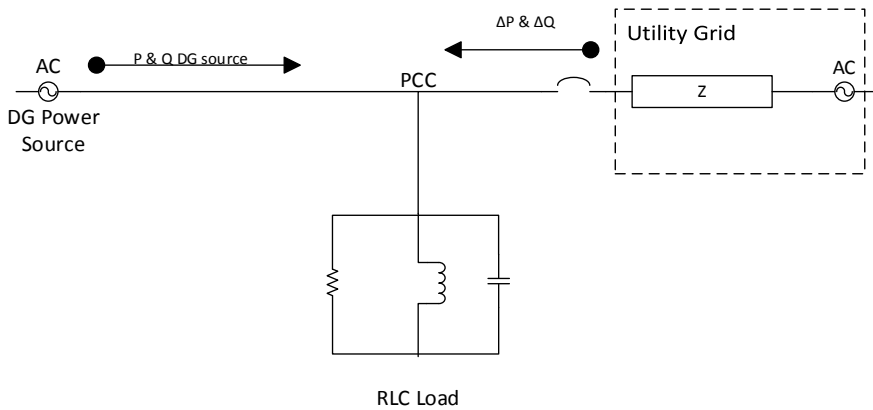


Fig. 6 DER, utility grid, and RLC load power circuit structural

Otherwise, if the system having lower threshold value, it showed that may be the system is having power quality indices [43].

Figure 6 shows the power circuit where utility grid, DER, and RLC load are connected at PCC. When islanding occurs, the voltage and frequency change depends on the power mismatch between DER and utility grid. During over or under voltage and frequency, the trip will be happened to the protection relay if the difference is more than threshold level. If local loads and DER output power are having small deviation, the over/under voltage and frequency methods not able to identify islanding correctly. Since, the DER monitors the terminal voltage, it may work individually without any assistant from other hardware. This shows that passive islanding method becomes one of the cheapest IDMs [49].

Table 2 shows comparison of few passive IDMs [48]. Each method has its own strength and weakness. For under/over voltage (UVP/OVP) and under/over frequency (UFP/OFP) methods, it is easy to implement, but the reaction time fluctuates. The weakness of (UVP/OVP) and (UFP/OFP) is that it has big NDZ. In phase jump detection (PJD) method, the islanding has been observed by sudden change of the phase angle, while monitoring the phase angle variation in terminal voltage and output current of the inverter. This method appropriates for inverter using a PLL in digital or analog form. Disadvantage of PJD is that it is difficult to implement because of the threshold, and it is unable to detect islanding condition if the DG power generation and power demand of local load are similar to each other. Total harmonic distortion (THD), using measurement of harmonic to detect the islanding in the system. If the THD is larger than a defined threshold, the inverter will automatically separate from the DGs and it considers that islanding occurs. The threshold selection and high Q factor detection are the issue arise for THD method. Voltage unbalance or VU differs due to the topology and small changes in the DG loading. VU islanding detection is only possible if there are unbalance in three-phase output and the DG

Table 2 Passive IDMs comparison [48]

Method	Implementation speed	Weakness	Improvement
UFP/OFP UVP/OVP	Implementation speed	Large NDZ	Compared using $P-V$ and $P-Q$ in constant current controlled inverters
PJD	Easy to implement but having random reaction time and variable	Islanding unable to identify if DERs equal to local load demand	
THD	Hard to implement and need to indicate threshold		Controlled by using a PLL
Voltage unbalance		Failed to detect islanding if low distortion of inverter or high quality load	

been monitored closely. The drawback of VU method is that it is not applicable for signal phase system.

3.3 Active Methods

Active methods depend on injecting perturbations in the distribution system to facilitate significant changes in the power system parameters and hence allow easy detection of the island [38, 50–52]. In active methods, the disturbances intentionally introducing to the utility grid and the reaction have been monitored to decide either the grid is still connected or not [29]. Active techniques have lesser NDZ, but their operation consequences in degrading the power quality because they introduce perturbations in the voltage and/or current at predefined intervals [53, 54].

Active methods are categorized into two main clusters, which are frequency shift and the voltage phase shift cluster. For frequency shift, one most popular technique is called slip mode frequency shift or SMS method. The frequency will be stabilized by the positive feedback once the main grid is detached. The positive feedback will cause inefficacy of the power quality which develop the transient response. Secondly called active frequency drift (AFD) method or frequency bias method, this method facing also has the same problem as SMS and has a bigger NDZ. Thirdly called Sandia frequency shift (SFS), major problem of this method is that it interrupts the power quality and system constancy. Fourthly called frequency jump detection (FJD) has the identical drawback as AFD, and it will lose efficiency in multiple inverter if the frequency is not synchronized. The multiple inverters terminate each other if the frequency is not synchronized and subsequent in detection failure. For three-phase systems, it can implement $D-Q$ implementation frequency feedback (DQIFF), this method has weaknesses in power quality and stability for the system. The random

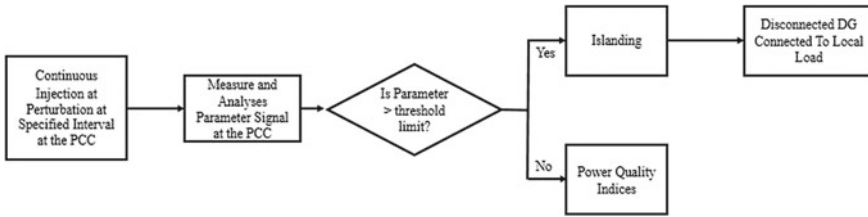


Fig. 7 Basic working principles of active islanding detection technique [43]

noise injection (RNI) is improper to apply in real world because it is hard to determine the magnitude of white noise, and therefore, onset points need to be identified.

Another group is under voltage phase shift techniques. First method is called Sandia voltage shift (SVS), and this method usually increases harmonic distortion in the system. Second popular method is differential level voltage shift (DVS), and this method is having less NDZ but takes a great level of distortion. High distortion will result in power degradation and noise in normal state. Active methods usually injecting negative sequence current and disturbance signals for islanding detection.

Mostly, active methods have small NDZ; however, they disturb the system. In this cause, it fails to work when more than one DERs are connected to the system if it has not been embedded with the identical active anti-islanding protection in the system [22-25].

Figure 7 shows the illustrations of basic working principle of active islanding detection. Firstly, continuous injection and perturbation at specified interval have been applied into the system. After the perturbation has been done, then local measurement of power system parameter at the PCC will be analyzed. If the parameter at the PCC is higher than threshold limit, islanding is detected and the DG is immediately disconnected from the load. But if the parameter is lesser than the threshold limit, this shows that the system is only suffering from the power quality indices.

3.4 Hybrid Methods

Hybrid methods are mixture of both active and passive schemes [55]. It introduces perturbations through active methods only if passive methods failed to detect islanding. Thus, this will reduce the number of perturbations injected into the system [56]. However, hybrid methods need longer time to detect the island compared to active and passive methods since system need to undergo two stages of steps if the passive method did not detect islanding in the system [38]. Throughout islanding detection, the passive method roles as principal, hence the active technique functions as a subordinate [43]. The mixture of passive and active methods will increase the efficiency of overall system. Figure 8 shows a simple working principles of hybrid IDMs [43].

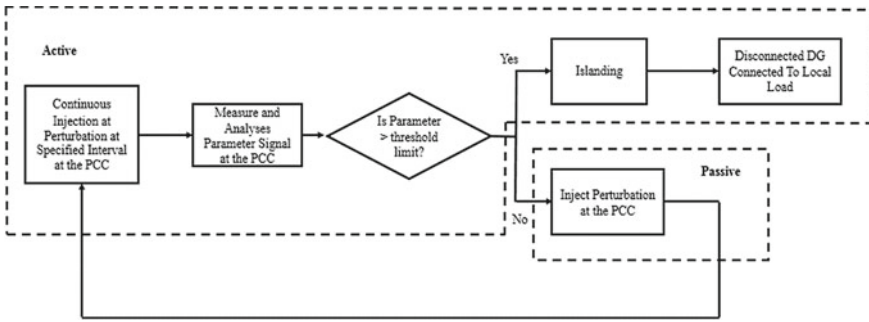


Fig. 8 Basic working principles of hybrid islanding detection technique

Table 3 Comparison between remote and local techniques

Characteristic	Remote techniques	Local techniques		
		Passive	Active	Hybrid
Detection time	Short	Short	Long	Long
Impact on power quality	None	None	High	Very small
NDZ	None	Large	Small	Very small
Consistency	Highly consistent	Less consistent	Highly consistent	Highly consistent
Cost	Expensive	Low	Low	Average

Advantages of the hybrid technique are small NDZ and improve the power quality of the system compared to passive and active techniques which work alone. However, this combination of two types of islanding techniques may double or triple the cost of the overall system. Since the process of islanding go through two different method, the islanding duration will take some time for detection.

Table 3 shows the summary comparison for remote and local techniques on the basis of detection time, the impact on power quality for each technique, NDZ, reliability, and the cost. There is a lot of research been done to upgrade the system such been done by [57], combine the changes between active power and reactive power.

3.5 Signal Processing Techniques

Signal processing upgrades the efficiency of passive IDMs. Signal processing will extract the hidden characteristic of measured signals to discover islanding [43]. Generally, Fig. 9 shows the procedure for islanding detection in signal processing techniques. All parameter signals such as voltage, current, frequency, and harmonics will be analyzed and detected disturbance by using a signal processing tool. Islanding

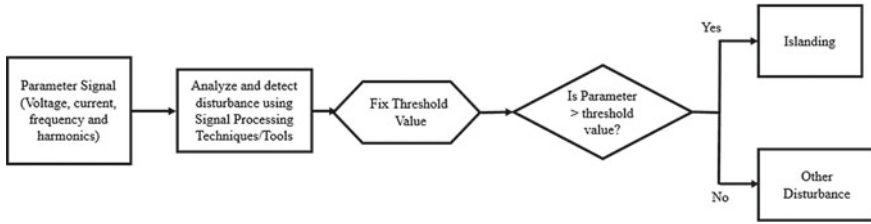


Fig. 9 Basic working principles of signal processing technique islanding detection

is detected if the reading value is bigger than the threshold value. Or else the parameter value is less than threshold value; it is defined as other disturbance in the system.

Figure 10 shows five popular signal processing tools which are used for IDMs. It includes Fourier transform (FT), wavelet transform (WT), S-transform (ST), hyperbolic S-transform (HST), and time-time transform (TTT).

Normally for frequency domain analysis, we called it as Fourier transform (FT). FT signifies a signal as a combination of sinusoidal form in variations of frequencies. FT capable of removing the stationary signal but incapable to detecting time distribution for variation of frequencies. It is also incapable to deal with transient information related with oscillation [58]. The modification of FT known as short time-Fourier transform (STFT) is the modification of Fourier transform. STFT is easy to achieve brief time-frequency characteristic, so it can give better performance

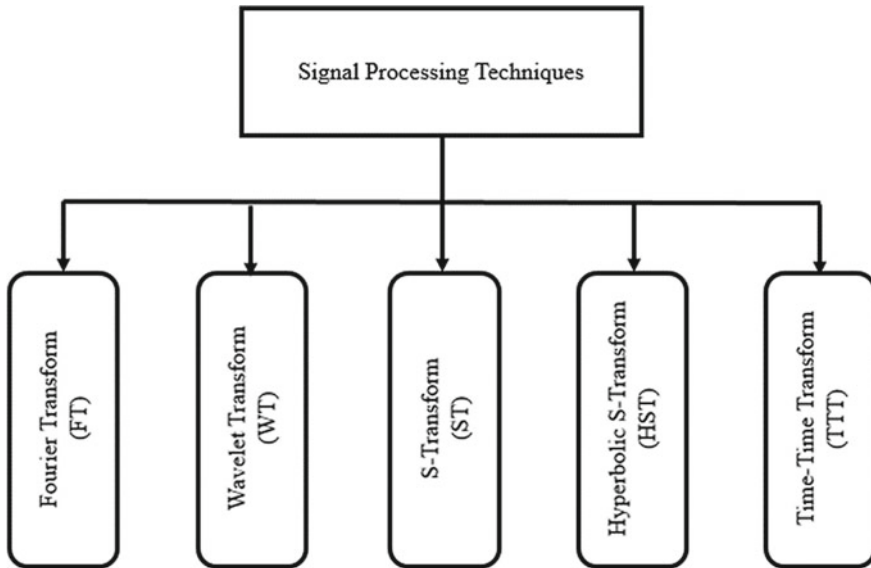


Fig. 10 General classification of signal processing techniques

compared to FT technique. However, STFT is failed to evaluate any static signal due to constraint in stationary frame [59].

Wavelet transform (WT) also one mutual tools for signal processing technique. It is based on the mathematical model using square integral with a same group theory as FT. WT symbolize as a signal in time and frequency domains [43]. This makes WT appropriate for observing the signals in time-frequency resolution; when necessary, it makes it among popular in power system applications. However, WT is oversensitive to noise, and it involves with computational complexity.

S-transform (ST) is the hybrid of STFT and WT properties. It is proposed to overcome weakness of WT which is batch processing, detection of various islanding, power quality (PQ) disturbance, and noise sensitivity [60]. ST is having better performance by detecting the disturbance signal efficiently and accurately. ST drawback, the islanding detection will be degraded if moving indicators are followed with transient.

The phasor information to WT is provided in hyperbolic S-transform (HST), for exposure of signatures in time–frequency plane. Since ST with Gaussian window failed to detect disturbance in time domain, HST with pseudo-Gaussian hyperbolic window has been introduced. HST is able to have greater control over window function, with enhanced timing at low or high frequency.

Time-time transform (TTT) using calculation of inverse Fourier Transform of discreet ST. Its mapping single- dimensional into two-dimensional signal in time domain. TTT do not require selection for different level and high-frequency component give more energy concentration for the signal to better localize in TTT.

Figure 11 [60] shows an islanding and PQ disturbance detection flowchart. The ST, HST, and TTT using the same step procedure. For all the transform methods, the first step is voltage signal acquisition. During this process, the transform components are computed to obtain HST or TTT matrix. Energy matrix and standard deviation will compute from the obtained matrix. If the energy matrix and standard deviation are higher than the fixed threshold, it means that the islanding occurs; if lower than the threshold, it means that there are power quality disturbance in the system such as swell, transient, notch, or THD.

According to previous discussion, we can determine that IDMs are having their advantages and disadvantages. Any how some major issues such as accuracy islanding detection, detection speed, and system compatibility in multiple and hybrid DERs environment is still need to be improved [43]. Success of the IDMs in distribution network came with how precise and fast of the islanding detection. However, each technique either remote, passive, or active is still suffering with lack in system which causes inaccuracies in discover islanding. Since signal processing and computational intelligence-based techniques are robust technique and able to dealing with complex nonlinear system, it becomes the best alternative for those issue [61].

In signal processing, wanted signal from input signal has been extracted and matches with the onset value. Selecting the range of onset value is uneasy job. If the onset value is in height, then islanding is unable to identify; but if the onset value is very low, it will trip the distributed generator even make the disturbance more badly [43]. In solving this issue, intelligent classifiers can be implemented.

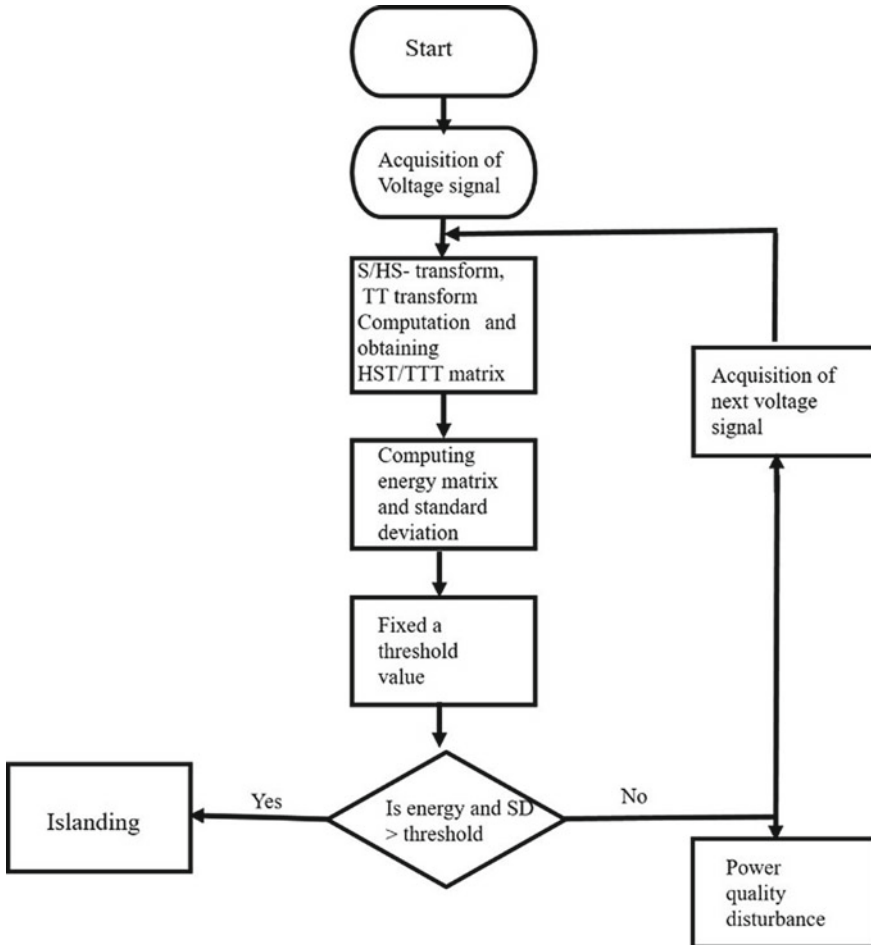


Fig. 11 Islanding and PQ disturbance flowchart

Furthermore, these techniques may be an appropriate choice to precisely discover the islanding phenomena in minimal time [61]. A technique, that applies copycat human intelligence, is called as computational intelligent technique. These techniques are shown in Fig. 12.

3.5.1 Artificial Neural Network (ANN)

Artificial neural network (ANN) is one of the popular method that has been implemented as a best solution for science and engineering problems. An ANN is a network which consists of nodes or neurons. An artificial neural network, called as neural network, is a mathematical model inspired by biological neural networks. A neural

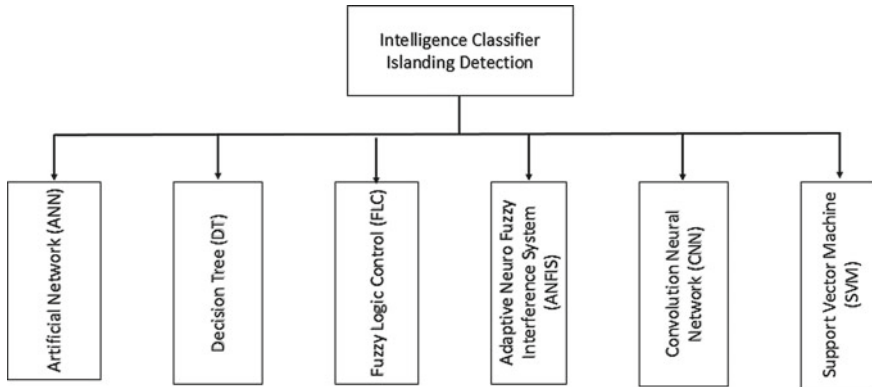
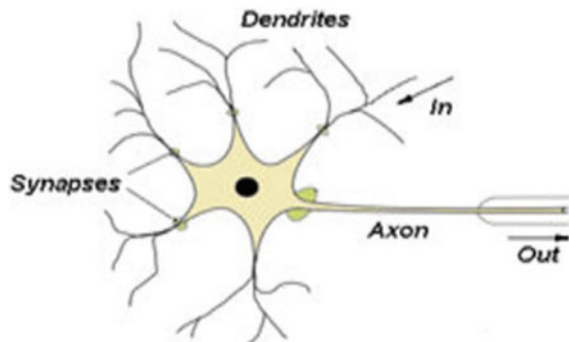


Fig. 12 General classification of intelligent classifier islanding detection technique

network contains of a connected set of artificial neurons, which processes an information using computation approach [62]. Neural network is having a characteristic and capabilities which is not available by other technology. ANNs are enormously parallel computing systems which include big number of modest processors with a lot of connections. One type of network sees as “artificial neurons” called as artificial neural networks (ANNs). Its function is to process information mainly in its related field. In engineering application, ANN is used for pattern recognition, forecast, and data compression. ANN basically consists of input (synapse) which multiplied with the weight. Weights assigned with each arrow represent the flow of the information. Mathematical function will be used to detect the activation of the neuron. The output is determined by another mathematical function. The neurons for the network is the sum of their inputs. The output should be the input neurons multiplied by a weight. Figure 13 shows the biological neuron. The output for specific input is determined by adjusting the weights. Some appropriate algorithms are implemented for adjusting the possible weight to get the wanted output from network. The process of adjusting

Fig. 13 Biological neuron [62]



the weight is called as learning or training process. The connection pattern of network can be classified into feed-forward networks and multilayer perceptron architecture.

In [59], the researcher used discrete wavelet transform to take out the data for islanding detection purpose. The data, which is been extracted, will then be feed to train the ANN to distinguish islanding and non-islanding events. A number of various type of islanding and non-islanding events need for the training. The discriminative topographies of voltage signals at the DER side are also taken during the training. The non-islanding cases include variations of load, capacitor load addition and disconnection of capacitive load, single line to ground fault, line-to-line fault, three phases to ground fault at different location of distribution network, and different DERs operating states. In conclusion, results of islanding detection show the maximum accuracy of 97.77%.

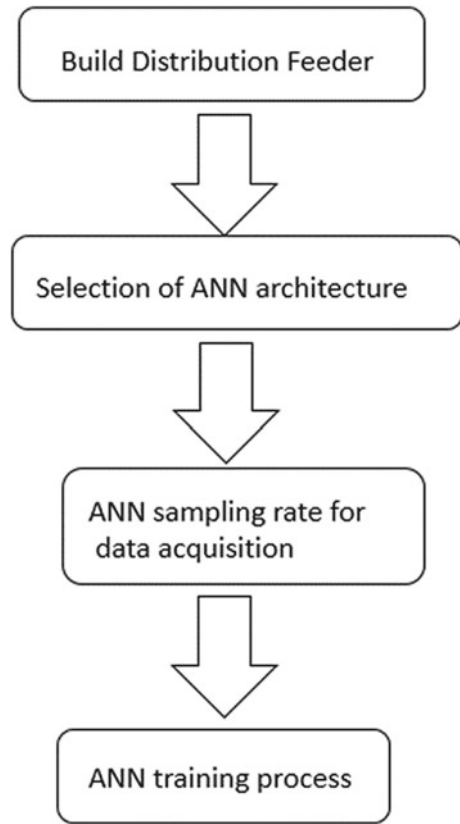
In [60], ANN-based hybrid IDMs which fit for synchronous-based DERs is suggested. Combination of the passive and active technique is driven to better dependability and precision of islanding detection. The author suggested six parameters in passive technique which are change of voltage for certain time (dv/dt), change of frequency over time (df/dt), change of frequency over change of power (df/dp), change of power over change of time (dp/dt), change of reactive power over change of time (dq/dt), and total harmonic distortion (THD) of current, while active technique uses the positive feedback of active/reactive power. The ANN is usually used with a signal processing technique such as wavelet transform since it is having an ability to remain the time domain properties and the frequency component that extract from the signal [63].

In [12], the transients generated during different types of disturbance in power system are analyzed using WT in order to determine an islanding. Combine three-phase signals into only one signal to observe the behavior of the system under various operating conditions. By combining the signal, it will reduce the computational burden for that system. The combination of the signal is called as "Modal Signal." The modal signal is processed using discrete WT in order to extract useful data for islanding detection. It is not easy to discover one suit criteria that differentiate islanding with 13 other events, especially when it is dealing with a feature which containing 12 variables. Due to this, ANN is used as a classifier to classify the occurrence of islanding event.

The hybrid ANN-based approach IDMs in distribution generation system is proposed in [64]. In this research, the passive and active technique have been used to increase the effectiveness and accurateness of the IDMs. Two different networks have been used for the test. First network showed the rate of false detection improved from 16.67 to 11.1% as compared to [65].

In [32], the new passive IDMs have been used for wind farm-based power generation system. The proposed method uses voltage and current measurements as parameters. This data were processed into Fourier transform for the second harmonic of voltage and current signals. Both signal, finally, will input to ANN for the islanding detection, refer Fig. 14 for the steps summary. This new IDM shows remarkable result, and it manages to identify islanding occurrence in very short time and efficiently till the load values (P&Q) not more or mirror to threshold value [32].

Fig. 14 A step using ANN concept



Another paper [66] takes ANN concept to develop the purposed method which used the step below in order to get the best outcome of their purposed method. The researcher used T circuit model by connecting $R-L$ series circuit and transformer for the distribution feeder. Distributed generator has been used as the DER for the distribution feeder. For the ANN architecture, the researcher proposed four layers of multilayer perceptrons. This ANN has been decided after few training and tests have been done using SimPowerSystems toolbox in MATLAB [66].

In order to determine the best sampling rate, four different samplings were tested. Two most sampling rates have been selected which are 64 and 128 samples/cycles since these both rates gave the best converge result while testing has been done. The next step is to train the ANN. While the training all the islanding situations, other disturbance need to be taken as a consideration. The researcher was having around 2000 simulations in the test system to get the final output for the system.

3.5.2 Decision Trees (DT)

Decision tree (DT) classifier is another type of pattern recognition tool. DT is able to offer a practical resolution for most possible inputs by considering their numerical discrepancies [48]. DT really is beneficial for the task impossible to be explained with logical approaches. Major benefit of using DT is that it is able to have short time training compared to other pattern recognition tools. Complex decision-making process can be easily been solved by decision tree algorithm with offers more straight forward resolutions which is easier to be understood [621].

A passive technique based on DT has been proposed in [18]. DT is having a huge set of data for training, and this causes a very hard work since the method has to train for each topology changing. This is causing a delay in islanding implementation. The wavelets transform as one of signal processing tools is introduced to extract voltage and current features and used a decision tree method to identify the islanding. Advantage of wavelet design in DT for island detection is that it makes the algorithm simpler and also lower computational effort [18].

DT-based classifiers perform a rectangular partitioning of the input space while the fuzzy models generate no axis parallel decision boundaries. In general, the initialization steps of the identification of the fuzzy model become very important. Grid-type partitioning and rule generation on extrema initialization result in complex and no interpretable initial models. To avoid such problems, a crisp decision tree, having high performance and computational efficiency, is proposed for initial partitioning of the input domain for the proposed fuzzy model [67]. In [67], the researcher used DT for initial classification boundaries. From the DT classification boundaries, the fuzzy membership functions (MFs) are created based on rule formulated for islanding detection. This proposed technique was tested with and without noise condition, and the outcome of this technique is the system able to detect 100% islanding [67].

Usually, the root nodes consist initial classification problem which divided into test parameter with the threshold or to set for further branch out. Each branch of the node will represent two decision (YES or NO), and it will continue branch out until it meet the match condition of the leaf node [68]. In one research, they used 11 features from 54 events to classify the islanding and non-islanding events.

3.5.3 Fuzzy Logic Control (FLC)

Fuzzy logic control (FLC) is one of the flexible tool which is able to deal with uncertain data. FLC is the best tool for a system which is not well defined by mathematical formulation. Main capability of a fuzzy logic system (FLS) is that it is able to express nonlinear input/output relationships by using if-then rules [69]. Comparing with DT classifier, FLS is having greater flexibility of the decision boundaries [5].

Three measurement parameters deal with FLC are voltage, ROCOF, and active power derivative (ROCOP) and is used in [48]. FLC also was presented from the conversion of DT, where the blend of fuzzy membership functions (MFs) and the rule to develop the fuzzy rule based [48]. Both proposed technique was simple for

implementation in online islanding detection. Plus it manages handle uncertainties or noise in the system. This proposed technique is working well at 00% detection rate tested with and without noise environment.

FLC is able to deal with complex systems such as MGs with different types of inaccurate inputs, variables, and disturbances such as renewable energy source consumed by varying and unpredictable load [17]. Fuzzy logic-based technique is having 11 parameters; change of frequency (df), change of voltage (dV), change of frequency over time(df/dt), change of voltage over time (dV/dt), change of power over time (dp/dt), change of frequency over power (df/dp), current time harmonic distortion (CTHD), voltage time harmonic distortion (VTHD), change of power factor ($Dp.f$), absolute voltage per power factor($V/p.f$), and gradient voltage per power factor ($V/p.f$) for islanding detection [61]. The simulation results showed 100% accuracy when using the fuzzy-based IDMs.

Fuzzy rule-based classifier technique for islanding detection in distributed generation has been presented in [67]. In it, the initial classification boundaries are using the decision tree (DT). From the DT classification boundaries, the fuzzy membership functions (MFs) are developed. The developed fuzzy rule-based classifier is tested using features with noise up to a signal-to-noise ratio of 20 dB and provides classification results without misdetection. The technique gives an efficient result of the islanding detection.

3.5.4 Adaptive Neuro-Fuzzy Interference System (ANFIS)

Adaptive neuro-fuzzy interface system or ANFIS is a very influential methodology for modeling nonlinear and complex systems since minimum input and output training data are needed. The compensations of ANN competence in learning processes and fuzzy logic control capability in managing ambiguous data is been combined in ANFIS. Hence ANFIS more capable of approaching nonlinear and uncertainties system without original mathematical equations. ANFIS method use the passive methods having less NDZ.

Three main properties of neuro-fuzzy network are the accurate learning, adaptive capabilities of the neural networks generalization, and fast learning capabilities of FLS. ANFIS plays a main roles while for modeling nonlinear and complex systems with less input and output training data. With minimum training data, the process will be faster and will contribute to good accuracy. Besides, ANFIS able to substitute the neural network topology together with FLS [69].

Generally, ANFIS function is like feed-forward back-propagation network. It used the future limits to determine forward function and use initial limits to determine backward function. Two learning techniques which appropriate for ANFIS are hybrid learning technique and back-propagation learning technique. Only zero or first order Sugeno inference system or approach learns the rules and membership functions from data taken into account in fuzzy rules. In conclusion, ANFIS function is to fine-tune the limits of a fuzzy system by applying a learning procedure using input-output training data [69].

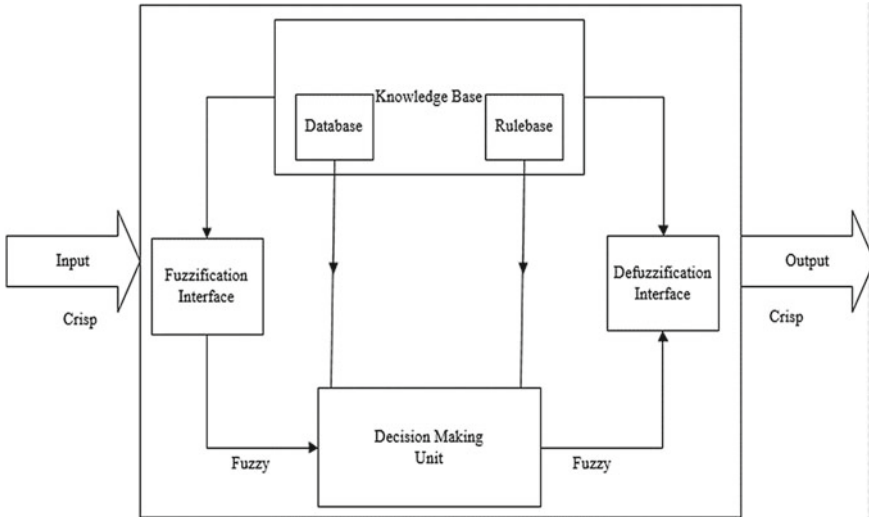


Fig. 15 Fuzzy inference system [68]

By referring Fig. 15, fuzzy inference systems are also known as a fuzzy rule-based system, fuzzy models, or fuzzy associative memory. Basically, a fuzzy inference system consists of four functional blocks which are (1) fuzzification interface, (2) knowledge base, (3) decision-making unit, and (4) defuzzification interface. In knowledge base, it includes i) rule base contains fuzzy if-then rules, and ii) data base consists of the membership functions of the fuzzy sets. A decision-making unit performs the inference operations on the rules. Fuzzification inference transforms the crisp inputs into degrees of match with linguistic values. Meanwhile, the defuzzification inference role is to transform the fuzzy results of the inference into a crumbly output. ANFIS is a combination of neural network (NN) and fuzzy logic (FL) components. Neural network is responsible to generate data for the fuzzy to modeling the logic solution. Fuzzy responsible to generate a suit membership function parameters that best allow the associated fuzzy inference system to track the given data as shown in Figs. 16 and 17.

In [69], the researcher using the ANFIS implementation based on passive method in wind turbine simulator. The outcome from this research showed that ANFIS-based algorithm detects the islanding more precise as compared to other islanding detection algorithm. Four different conditions have been observed for this paper which is match power condition, mismatch power condition, motor starting condition, and capacitor bank switching condition. All these four conditions are very important for validation of the ANFIS algorithm.

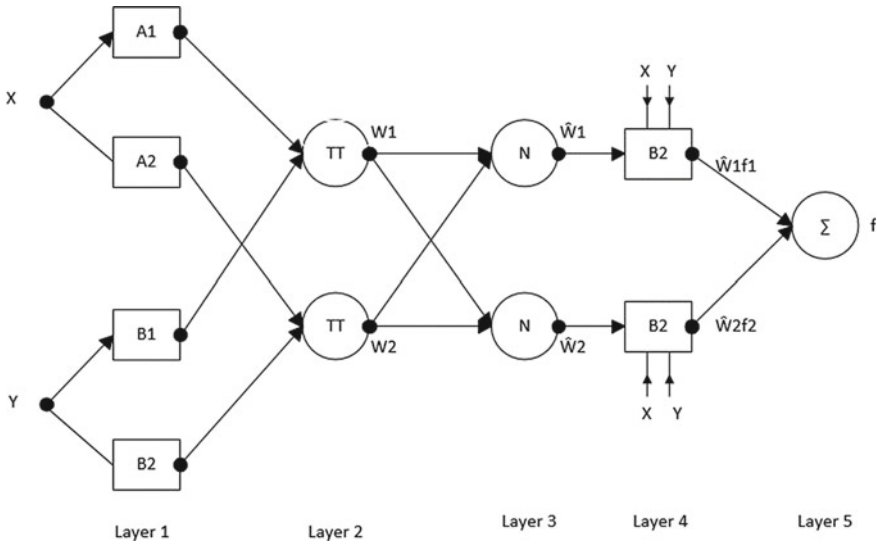


Fig. 16 ANFIS architecture [68]

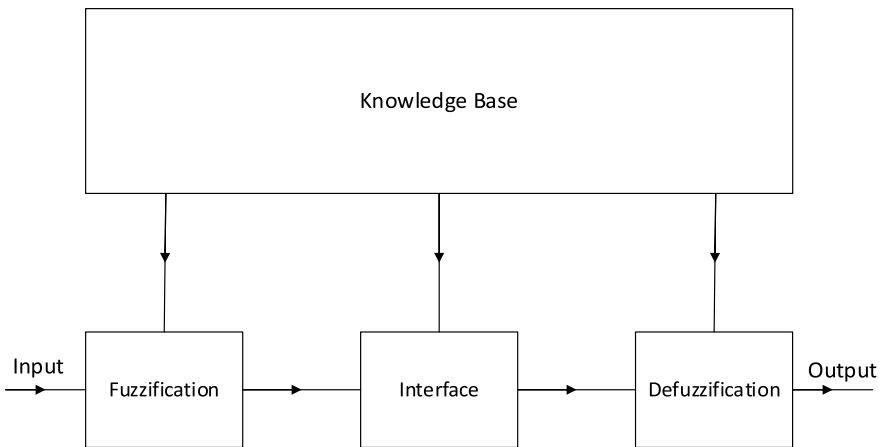


Fig. 17 Adaptive neuro-fuzzy inference system [5]

3.5.5 Convolution Neural Network (CNN)

Convolution neural network (CNN) is one of the latest identification classifier neural network. CNN usually use image classification based for getting the result. The biggest challenge for the neural network system is to have huge label data to train the model [70]. In [70], the researcher discussed new islanding detection method which is based on the image classification with CNN. In this paper, the time series data is

converted into images. Those time series data includes the information regarding the distinct patterns for grid connected and its island possible islanded system. Series data is converted into image by using wavelet transform.

In this paper, time energy frequency density which converted using wavelet transform is called scalogram. The scalogram image for the time series data is generated by applying continuous wavelet transform (CWT). Since CNN is also called as supervised machine learning, huge number is required for the training and testing purpose. The purpose technique is used a 100 kW grid connected PV system in order to create the data set. The researcher used MATLAB/Simulink software to simulate the system. In total, there were 104 cases of islanding events and 101 of non-islanding events which been used as a data set.

A lot of similarities can be seen between CNN and neural network (NN). Basically, both are having neurons, biases, and learning weight. In CNN, there have more than one layers to process, classify, and segment the data to make a better prediction of its outcome.

Figure above shown the general architecture for CNN. In CNN, there are three main layers called input image, image to feature, and feature to classification layer. In first layer, the system will be feed with all the data sets. In the second layer, all the scalogram image will go through few steps such as convolution, rectified linear unit, and the max polling. Convolution is known as the main layer in CNN. In this layer, the image data will be filter by a learning filter which known as kernel. The image will be convoluted few times using several of kernel. Load of kernel images with different features will be used as an input for the next layer. Output from the convolution normally is linear in nature. This output will need to go through the activation function to achieve nonlinear transformation. Beside activation function, pooling layer is also one of the important layer in the convolution. Pooling layer will do the down sampling, and the features have been extracted from the convolution; it also reduces the dimension; furthermore, softmax layer takes n-dimension input vector to calculate all possibilities of features. In fully connected layer, all the feature will be listed in order to do the classification process. This classification process will be classified to the cases of islanding and non-islanding [70].

3.5.6 Support Vector Machine (SVM)

Support vector machine or SVM is also one of the intelligent islanding method. Two major issues which encourages the researcher to use this approach was the large NDZ of the passive method and the threshold setting requirements which available in most islanding conventional method [44].

SVM utilizes statistical learning technique which provides better precision in high-dimensional features spaces. Pattern classification is based on structural risk minimization method [60]. SVM is one innovative and capable methods for learning and separating function in pattern recognition (classification) or performing function estimation in the regression problem. SVM originates from the statistical learning theory (SLT), developed by Vapnik for a “distribution free learning from data” [71].

In research by [54], the current and voltage are measured at various location of the microgrid. Wavelet transform is used to extract the number of available features from the decomposed voltage and current signal. Based on the extracted features, the SVM classifier is trained to detect islanding and non-islanding conditions in the microgrid [44].

In [7] passive IDM, a multiple feature-based SVM classification technique is proposed. It used five network parameters: frequency, voltage, rotor angle, rate of change of voltage, and rate of change of frequency. This method tested large number of islanding and no-islanding events. The trained SVM is tested on typical distribution network containing multiple DG units. The results proved that the proposed technique works effectively with high degree of accuracy. Under different network possibilities and critical level of power imbalance that may happen during islanding.

In [71], the author using IEEE 9 bus system with an islanding and non-islanding is been detected by using SVM classifier training and testing in MATLAB environment. The performance of SVM is better than ANN because ANN basically follow empirical risk minimization (ERM) where the SVM based on structural risk minimization (SRM) [71]. In this SVM approach, it tries to find the hyperplane to separate data point according to their classes. Hyperplane should be optimal hyperplane, and the separation classes should be maximum. The advantage of using SVM approach is that it can minimize the dimension which a bulk of dimension exist in conventional pattern recognition system.

4 Discussion

The characteristics of the IDMs have been summarized in Table 4. Remote method is the most reliable and gives no impact on power quality. However, it is the most expensive IDMs. The passive method also has no impact in power quality and will detect the islanding promptly. However, it fails to work in closely mismatched conditions. Therefore, the NDZ for passive islanding is larger compared to active islanding method. Active method detects the islanding in an imperfect match between generation and demand. This method will reduce error detection ratio. Since the active method introduces the perturbation in the system, this will cause slow detection time because extra time is needed for perturbation to be observe in the system. The perturbation usually degrades the power quality and also the system stability. The combination of passive and active method is known as hybrid method. This method gives small NDZ. Since hybrid method is a combination of two methods, the system becomes complex. The combination can improve multiple performance indices simultaneously by their compensation.

Since the conventional IDMs facing a lot of lacking in the system, the new era computational intelligence-based technique for IDMs has been introduced. This technique is one of the best choice since it may detect the islanding accurately in very minimum time [72] (Fig. 18).

Table 4 IDMs comparison table [3, 43]

Characteristic	Remote techniques	Local techniques			Signal processing-based methods	Intelligence classifier-based methods
		Passive	Active	Hybrid		
Principle	Communication between utility grid and DERs	Monitor natural effects of islanding	Inject noise	Combine passive and active	Monitor changes in system parameter	Applied to AI
Detection time	Small	Small	Long	Long	Very small	Very small
Impact on power quality	No	No	High	Very small	No	No
NDZ	None	Large	Small	Very small	No	No
Reliability	High	Low	High	High	Very high	Very high
Cost	High	Low	Low	Low	Low	Low
Effect on distribution System	None	None	Direct influence	Lower than active	None	None
Multiple DERs operation	Promising	Promising	Not promising	Not promising	Promising	Promising

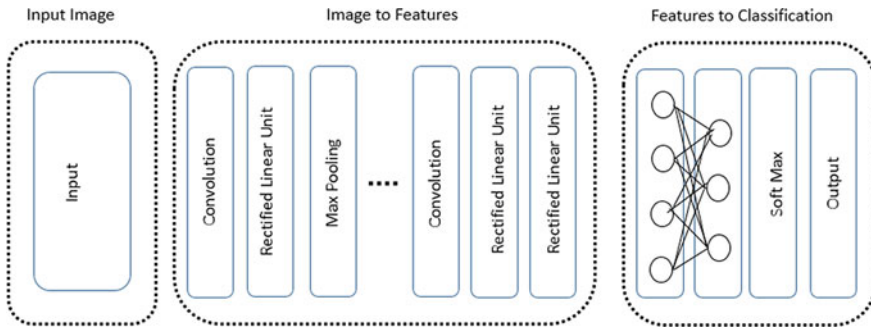


Fig. 18 General architecture for CNN [70]

Figure 19 showed that the most utilized computational intelligence-based techniques used for islanding detection are ANN and DT method, secondly FLC method, and thirdly is system ANFIS-based method followed by PSO and support vector machine (SVM) method. The least preferred method is artificial immune system (AIS). In AIS, islanding detection is based on two modules which is T-module and B-module. T-module is used to detect the islanding condition and B-module to improve the detection coverage space [45, 73]. ANN and DT methods become most popular intelligence technique applied because it gives more than 96% accuracy beside both easy to implement. The percentage accuracies of these techniques have been presented in Table 5.

DT-based classifiers execute a rectangular partitioning of the input space, while the fuzzy models generate no axis parallel decision boundaries. Main improvement of fuzzy rule-based classifiers over crisp DTs is that it is having superior flexibility of the decision boundaries. It is much easier to interpret and understand fuzzy classifiers compared to DT classifiers. The initialization steps of the identification of the fuzzy model are very important in DT. Common problem is arised such as grid-type

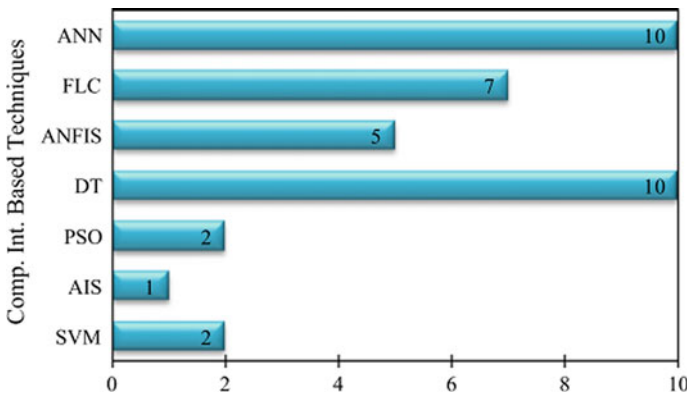


Fig. 19 Computational intelligence-based technique used for islanding detection [61]

Table 5 Accuracy percentage of IDMs

IDMs	Accuracy (%)	IDMs	Accuracy (%)
Over/under voltage	78.81	SOM neural network	97.92
Over/under frequency	90.24	SOM neural network	98.19
Voltage vector shift (VVS)	74.05	PNN	90
ROCOF-based technique	93.81	MPNN	97.4
DT classifier-based relay	96.43	Fuzzy and DT-based technique	100
Intelligent-based relay	83.33	ANFIS-based technique	100
Voltage-based DT relay	98	DT and DWT-based technique	96.43
ANN-based technique	97.77	DT-based relay	99.61
ANN-based technique	99.1	SVM-based technique	78
ANN-based technique	88.9	SVM-based technique	98.2

partitioning, rule generation on extrema initialization, result in complex, and no interpretable initial models. In order to eliminate such problems, a crisp decision tree, high performance, and computational efficiency are proposed for initial partitioning of the input domain for the best proposed fuzzy model [67].

To find the new IDMs is still going on. The focus is not only on the local methods which are active, passive and hybrid methods. But more focus has been made on intelligent classifier islanding detection method due to their simplicity, robustness, and higher accuracy. The modification on the intelligent classifier is also growing exponentially. Since intelligent classifier is giving the best solution and highest efficiency with fast detection, it is perfect for islanding detection.

A number of researches are still focusing on the passive methods due to simplicity in implementation and inexpensiveness [31] and [26]. Most of the papers have been published for islanding detection technique topic using the simulation software such as MATLAB simulation, PS-CAD, and RSCAD. Most of the simulations prefer to use more than one of distributed energy resources to understand the behavior of the power system.

5 Conclusions

This paper presents a review on intelligent classifier techniques for IDMs in distribution network. From all the literature, review showed that the application of intelligent classifier combination with signal a processing tool contributes to the fastest and reliable IDMs. There is no negative and bad consequences that have been discovered for this combination.

Combination of signal processing tools and intelligent classifier overcomes the drawback of the threshold setting and high NDZ if passive method works alone. Efficient intelligent classifier IDMs are very crucial for a successful islanding operation in microgrid and distribution network. Intelligence classifier improves the consistency and power quality of over power system. Thus, intelligent classifier-based

techniques have higher efficiency, more than 90% have a great potential to be used in real-time application.

Acknowledgements The authors would like to thank the Ministry of Higher Education, Malaysia, under Large Research Grant Scheme (LRGS): LRGS/1/2019/UKM/01/6/3 and University of Malaya, Malaysia, for providing financial support under the research grant RU-SATU Joint Research 2020: ST009-2020.

References

1. D. Archer, Global warming, in *Understanding the Forecast* (Blackwell Publishing, Ltd., Malden, MA, 2007)
2. N.H. Stern, *The Economics of Climate Change: The Stern Review* (Cambridge University Press, 2007)
3. Y. Hamdaoui, A. Maach, Smart Islanding in smart grids, in *Smart Energy Grid Engineering (SEGE)* (IEEE, 2016), pp. 175–180
4. J. Notenboom, P. Boot, *An Essay on the Colorful Scene of Europe's Energy Transition* (2016)
5. S. Kumar, G. Tyagi, M. Tyagi, Islanding detection in distributed generation by neuro-fuzzy approach. *Int. J. Emerg. Technol. Adv. Eng.* **4**, 52–57 (2014)
6. A. Datta, A. Ray, D. Mukherjee, H. Saha, Selection of islanding detection methods based on multi-criteria decision analysis for grid-connected photovoltaic system applications, in *Sustainable Energy Technologies and Assessments-Science Direct*, 2014
7. A. Khamis, H. Shareef, A. Mohamed, E. Bizkevelci, Islanding detection in a distributed generation integrated power system using phase space technique and probabilistic neural network. *Neurocomput.-Sci. Direct* **148**, 587–599 (2015)
8. A. Kaur, A.G. Sinha, *Islanding Detection And Protection* (2014)
9. C.S. Chandrakar, B. Dewani, D. Chandrakar, An assessment of distributed generation islanding detection methods. *Int. J. Adv. Eng. Technol.* **5**, 218–226 (2012)
10. C. Li, Y. Cao, Y. Kuang, L. Zeng, B. Fang, A review of islanding detection methods for microgrid. *Renew. Sustain. Energy Rev.* **35**, 8 (2014)
11. M.A. Eshraghi, M.S. Sariri, V. Schwarzer, R. Ghorbani, D. Block, F.S.E. Center, Islanding detection and over voltage mitigation using wireless sensor networks and electric vehicle charging stations (2016)
12. M.S. ElNozahy, E.F. El-Saadany, M.M. Salama, A robust wavelet-ANN based technique for islanding detection, in *Power and Energy Society General Meeting, 2011* (IEEE, 2011), pp. 1–8
13. O.N. Faqhruldin, E.F. El-Saadany, H.H. Zeineldin, A universal islanding detection technique for distributed generation using pattern recognition. *IEEE Trans. Smart Grid* **5**, 1985–1992 (2014)
14. A. Salam, A. Mohamed, M. Hannan, Technical challenges on microgrids. *ARPN J. Eng. Appl. Sci.* **3**, 64–69 (2008)
15. F. Yang, N. Xia, Q.-L. Han, Event-based networked Islanding detection for distributed solar PV generation systems. *IEEE Trans. Industr. Inf.* **13**, 322–329 (2017)
16. N. Lidula, A. Rajapakse, Microgrids research: a review of experimental microgrids and test systems. *Renew. Sustain. Energy Rev.* **15**, 186–202 (2011)
17. R. Al Badwawi, W. Issa, T. Mallick, M. Abusara, Power management of AC islanded microgrids using fuzzy logic (2016)
18. L. Gao, J. Liu, A review of Islanding detection methods for photovoltaic generation system. *Int. J. Multim. Ubiquitous Eng.* **11**, 1–12 (2016)
19. W. El Khatam, M.M.A Salama, Distributed generation technologies, definition and benefits. *Electric Power Syst. Res.* **71**, 119–128 (2004)

20. A. Khaledian, M.A. Golkar, A new power sharing control method for stability enhancement of islanding microgrids, in *2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)*, 2016, pp. 1–5
21. H. Jiayi, J. Chuanwen, X. Rong, A review on distributed energy resources and MicroGrid. *Renew. Sustain. Energy Rev.* **12**, 2472–2483 (2008)
22. E. Planas, A. Gil-de-Muro, J. Andreu, I. Kortabarria, I.M. de Alegria, General aspects, hierarchical controls and droop methods in microgrids: a review. *Renew. Sustain. Energy Rev.* **17**, 147–159 (2013)
23. Q. Lei, F.Z. Peng, I.J. Balaguer, Islanding control of DG in microgrids, in *IEEE 6th International Power Electronics and Motion Control Conference, 2009. IPEMC'09, 2009*, pp. 450–455
24. G. Antonova, M. Nardi, A. Scott, M. Pesin, Distributed generation and its impact on power grids and microgrids protection, in *2012 65th Annual Conference for Protective Relay Engineers, 2012*, pp. 152–161
25. J.A. Laghari, H. Mokhlis, A.H.A. Bakar, M. Karimi, A new islanding detection technique for multiple mini hydro based on rate of change of reactive power and load connecting strategy. *Energy Convers. Manag.*, pp. 215–224 (2013)
26. A.F. Sagar, M. Shamshiri, C.K. Gan, R.T. Daram, M.R. Ab Ghani, Reactive power imbalance method for Islanding detection in micro-grid operation. *Int. J. Recent Technol. Eng. (IJRTE)* **2**, 5 (2014)
27. N. Ghadimi, H. Ebrahimian, R. Ghadimi, A. Danandeh, A new active method to diagnose the anti-islanding mode protection for photovoltaic system. *World Appl. Sci. J.* **15**, 1593–1597 (2011)
28. H. Vahedi, G.B. Gharehpetian, M. Karrari, Application of duffing oscillators for passive islanding detection of inverter-based distributed generation units. *IEE Trans. Power Deliv.* **27**, 1973–1983 (2012)
29. W.G.M.A.S. Aljankawey, L. Chang, C.P. Diduch, Passive method-based islanding detection of renewable-based distributed generation: the issues, in presented at the *IEEE Electrical Power & Energy Conference*, 2010
30. S. Mekhilef, N.A. Rahim, Implementation of grid-connected photovoltaic system with power factor control and islanding detection, in *35th Annual IEEE Power Electronics Specialists Conference*, pp. 1409–1412 (2004)
31. S. Raza, H. Mokhlis, H. Arof, J.A. Laghari, H. Mohamad, A sensitivity analysis of different power system parameters on islanding detection. *IEEE Trans. Sustain. Energy* **7**, 461–470 (2016)
32. A.G. Abd-Elkader, D.F. Allam, E. Tageldin, Islanding detection method for DFIG wind turbines using artificial neural networks. *Electric. Power Energy Syst.—Sci. Direct* **62**, 335–343 (2014)
33. N. Strath, *Islanding Detection in Power Systems* (Department of Industrial Electrical Engineering and Automation, Lund University, 2005)
34. M. Hashemi, H. Mahdian, A.A. Ghadimi, A new method for islanding detection the grid connected inverters in case of unbalanced loads. *Indian J. Sci. Technol.* **6**, 12 (2013)
35. D.-U. Kim, S. Kim, Anti-islanding detection method using phase-shifted feed-forward voltage in grid-connected inverter. *IEEE Access* **7**, 147179–147190 (2019)
36. R. Walling, N. Miller, Distributed generation islanding-implications on power system dynamic performance, in *Power Engineering Society Summer Meeting, 2002 (IEEE, 2002)*, pp. 92–96
37. M.H. Moradi, M. Eskandari, P. Siano, Safe Transition from Connection Mode to Islanding Mode in Microgrids, in *2016 24th Iranian Conference on Electrical Engineering (ICEE), 2016*, pp. 1902–1907
38. S. Talwar, *Islanding Detection In Distribution System Embedded with Renewable Based Distributed Generation* (B. Eng, Graduate Academic Unit of Electrical and Computer Engineering, University of Ontario Institute of Technology, Canada, 2012)
39. H. Jouybari-Moghaddam, S. Hosseinian, B. Vahidi, Active distribution networks islanding issues: an introduction, in *2012 11th International Conference on 2012 Environment and Electrical Engineering (EEEIC)*, pp. 719–724

40. Z. Yipeng, T. Yun, R. Jiayu, Y. Libin, L. C. X. Chao, Micro-grid islanding detection based on PQ active method, in *35th Chinese Control Conference*, Chengdu, China, 2016, p. 5
41. K. Colombage, J. Wang, C. Gould, C. Liu, PWM harmonic signature based islanding detection for a single-phase inverter with PWM frequency hopping. *IEEE Trans. Ind. Appl.* **53**, 411–419 (2017)
42. I. Mazhari, L. Beghou, J. Enslin, B. Parkhideh, S. Bhowmik, Locking frequency band exposure method for islanding detection and prevention in distributed generation, in *2014 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2014, pp. 4361–4366
43. S. Raza, H. Mokhlis, H. Arof, J. Laghari, L. Wang, Application of signal processing techniques for islanding detection of distributed generation in distribution network: a review. *Energy Convers. Manag.* **96**, 613–624 (2015)
44. G.S. Rao, G.K. Rao, SVM based pattern recognised islanding detection approach in a multiple distributed generation system. *Int. J. Comput. Technol. Appl.* **10**, 189–196 (2017)
45. A. Hatata, El-H. Abd-Raboh, B.E. Sedhom, A Review of anti-islanding protection methods for renewable distributed generation systems. *J. Electr. Eng.*
46. X. Guoa, D. Xub, B. Wub, Overview of anti-islanding US patents for grid-connected inverters. *Renew. Sustain. Energy Rev.* 311–317 (2014)
47. R.A.L.T. Daram, *Islanding Detection Using Passive Reactive Power Imbalance Method* (Bachelor in Electrical Engineering (Industrial Power), Faculty of Electrical Engineering, Universiti Teknikal Malaysia Melaka, 2013)
48. A. Khamis, H. Shareef, E. Bizkevelci, T. Khatib, A review of islanding detection techniques for renewable distributed generation systems. *Renew. Sustain. Energy Rev.* **28**, 483–493 (2013)
49. R. Walling, Application of direct transfer trip for prevention of DG islanding, in *Power and Energy Society General Meeting, 2011* (IEEE, 2011), pp. 1–3
50. S. Liu, S. Zhuang, Q. Xu, J. Xiao, Improved voltage shift islanding detection method for multi-inverter grid-connected photovoltaic systems. *IET Gener. Transm. Distrib.* **10**, 3163–3169 (2016)
51. T. Skocil, O. Gomis-Bellmunt, D. Montesinos-Miracle, S. Galceran-Arellano, J. Rull-Duran, Passive and active methods of islanding for PV systems, in *13th European Conference on Power Electronics and Applications, 2009. EPE'09*, pp. 1–10
52. H.-L. Jou, W.-J. Chiang, J.-C. Wu, Virtual inductor-based islanding detection method for grid-connected power inverter of distributed power generation system. *IET Renew. Power Gener.* **1**, 175–181 (2007)
53. M. Vatani, M. Sanjari, G. Gharehpetian, Islanding detection in distribution networks with distributed energy resources before main circuit breaker opening. *Univ. J. Electr. Electron. Eng.* **2**, 1–5 (2014)
54. M.M. Hanif, M. Basu, K. Gaughan, A discussion of anti-islanding protection schemes incorporated in a inverter based DG (2011)
55. M.S.M.I.H.M.Q. Ahsan, Impact of sizes of islands on the stability of a faulted power system, presented at the *2nd International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT) 2015*, Iahangirnagar University, Dhaka-I 342, Bangladesh, 2015
56. L.P. Raghav, T. Sandhya, An active frequency drift method for an islanding detection of grid connected microturbine generation system. *Int. J. Innov. Res. Sci., Eng. Technol.* **3**, 111–117 (2014)
57. U.K. Jhuma, S. Mekhilef, M. Mubin, S. Ahmad, M. Rawa, Y. Alturki, Hybrid islanding detection technique for Malaysian power distribution system, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, 2020, pp. 785–790
58. M. Karimi, H. Mokhtari, M.R. Iravani, Wavelet based on-line disturbance detection for power quality applications. *IEEE Trans. Power Delivery* **15**, 1212–1220 (2000)
59. Y.H. Gu, M.H. Bollen, Time-frequency and time-scale domain analysis of voltage disturbances. *IEEE Trans. Power Deliv.* **15**, 1279–1284 (2000)
60. S.R. Mohanty, N. Kishor, P.K. Ray, J.P. Catalo, Comparative study of advanced signal processing techniques for islanding detection in a hybrid distributed generation system. *IEEE Trans. Sustain. Energy* **6**, 122–131 (2015)

61. J.A. Laghari, H. Mokhlis, M. Karimi, A.H.A. Bakar, H. Mohamad, Computational intelligence based techniques for islanding detection of distributed generation in distribution network: a review. *Energy Convers. Manag.* 139–152 (2014)
62. N. Gupta, Artificial neural network. *Netw. Complex Syst.* 3, 24–28 (2013)
63. Y. Fayyad, A. Osman, Neuro-wavelet based islanding detection technique, in *Electric Power and Energy Conference (EPEC), 2010 IEEE*, 2010, pp. 1–6
64. R. Ghazi, N. Lotfi, A new hybrid intelligent based approach to islanding detection in distributed generation, in *Universities Power Engineering Conference (UPEC), 2010 45th International*, 2010, pp. 1–5
65. K. El-Arroudi, G. Joos, I. Kamwa, D.T. McGillis, Intelligent-based approach to islanding detection in distributed generation. *IEEE Trans. Power Deliv.* 22, 828–835 (2007)
66. V. Merlin, R. Santos, A. Grilo, J. Vieira, D. Coury, M. Oleskovicz, A new artificial neural network based method for islanding detection of distributed generators. *Int. J. Electr. Power Energy Syst.* 75, 139–151 (2016)
67. S. Samantaray, K. El-Arroudi, G. Joos, I. Kamwa, A fuzzy rule-based approach for islanding detection in distributed generation. *IEEE Trans. Power Deliv.* 25, 1427–1433 (2010)
68. S. Vyas, R. Kumar, R. Kavasseri, Data analytics and computational methods for anti-islanding of renewable energy based distributed generators in power grids. *Renew. Sustain. Energy Rev.* 69, 493–502 (2017)
69. N. Ghadimi, B. Sobhani, Adaptive neuro-fuzzy inference system (ANFIS) islanding detection based on wind turbine simulator. *Int. J. Phys. Sci.* 8, 1424–1436 (2013)
70. S.K.G. Manikonda, D.N. Gaonka, *An Islanding Detection Method Based on Image Classification with Convolution Neural Network* (Institute of Engineering Technology, 2019)
71. S. Savitha, N. Kayalvizhi, The support vector machine technique for islanding detection in distributed generation. *Int. J. Adv. Res. Electr., Electron. Instrum. Eng.* 5, 7 (2016)
72. J. Laghari, H. Mokhlis, M. Karimi, A. Bakar, H. Mohamad, Computational Intelligence based techniques for islanding detection of distributed generation in distribution network: a review. *Energy Convers. Manag.* 88, 139–152 (2014)
73. G. Yin, A distributed generation islanding detection method based on artificial immune system, in *Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES*, 2005, pp. 1–4

Evaluating Classical and ANN-Based Load Forecasting Techniques Using Univariate and Multivariate Analysis



Kuheli Goswami, Ayandeep Ganguly, Nayan Manna,
and Arindam Kumar Sil

Abstract The consumption of electricity per capita is an indicator of a country's growth. It acts as a comprehensive reflection of the industrial growth of a country. The delivery of uninterrupted power of good quality is required to ensure rapid growth of the country's infrastructure. A judicious and proper planning is very essential, which helps in purchasing of energy, generation and infrastructure development, but also helps in load management. So, this makes the advance forecasting of load a very essential part of demand side management with the objective of renewable energy integration. Numerous methods are available for load forecasting and specifically the forecasting of peak load. Long short-term memory (LSTM) network and artificial neural network (ANN) are two different applied machine learning algorithms in this area. In recent days, autoregression integrated moving average with exogenous variables technique (ARIMAX) has proved itself a tool successful for prediction of future demand data. In this paper, an attempt has been made to compare these methods and determine their advantages and disadvantages. With an objective to find the best application amongst these methods which gives optimum results in load forecasting, a case study has been done. This paper has been able to show that ARIMAX is performing better than ANN and LSTM as future load forecasting tool.

Keywords Load forecasting · Artificial neural network · Long short-term memory network · Autoregressive integrated moving average with exogenous variables · Mean absolute error

K. Goswami (✉)

Brainware Group of Institutions-SDET, Kolkata 700125, India

A. Ganguly · N. Manna

Haldia Institute of Technology, Haldia 721657, India

A. K. Sil

Jadavpur University, Kolkata 700032, India

e-mail: arindamkumar.sil@jadavpuruniversity.in

1 Introduction

Nowadays, efficient energy consumption is one of the greatest concerns. There are several factors related to climate such as temperature, seasons and humidity as well as some sociological factors such as week days, weekends and holidays which govern the daily demand profile. These daily variations result in increasing the complexity of the series which exhibits several levels of seasonality. An efficient forecasting tool for electrical load is very important to reduce the stress on grid by the renewable energy integration. It can be utilized effectively to get knowledge about the present and future load demand scenario. Thus, load forecasting is an indispensable part of power system planning. A high accuracy model is extremely important for management of peak load. The significant impact of STLF on the economic operations of different utilities lies in the advance information it provides, which can help in economic generation scheduling and determine the available transmission capacity transactions.

2 Review of Load Forecasting Methods

To supply reliable and quality power, a system should have the capability to match the demand with the present generation dynamically with time, within the existing network. Forecasting of load helps the power to anticipate the amount of power demand and to make important decisions. The classification of load forecasting based on the forecasting horizon can be made as short-term load forecasting (STLF), medium-term load forecasting (MTLF) as well as long-term load forecasting (LTLF). Different methods of load forecasting techniques have been reviewed by Srivastava et al. [1]. In an article, the author has proposed an ARIMA model for load forecasting [2]. A real-time load forecasting has been proposed by an author [3]. ANN applications in load forecasting have been acknowledged by the researcher in recent times [4–8]. Recently, few works have been done for short-term load forecasting and peak load forecasting using LSTM which is a special kind of recurrent neural network [9–11]. A load model in hourly basis using time series model has been proposed by the author [12]. Univariate and multivariate models for daily as well as monthly snow water forecasting, respectively, for Ontario, Canada, were proposed by the authors in a work [13]. Weather condition has a huge impact in load forecasting, and this was detailed in an article [14]. In a research article, an author has checked the accuracy of ARIMAX and Box–Jenkins-based load forecasting models [15]. An artificial neural network-based day ahead forecasting model was proposed by Sahay et al. [16] for PJM and New England Electricity Market. It is known as a fact that day ahead forecasting gives intended results. But for week ahead forecasting, the accuracy of the load forecasting models needs to be improved. So, the problem has been addressed here using neural network, time series and LSTM for the electricity market of the state of West Bengal.

3 Mathematical Model Design

Here, in this paper, design of mathematical model using ARIMA, ARIMAX, LSTM and ANN has been discussed.

3.1 ARIMA

Autoregressive integrated moving average is a univariate time series method. ARIMA and neural networks are defined as statistically sophisticated and mathematically complex methods.

Using the given flow chart, the approach for designing an ARIMA model for forecasting is explained briefly [9–11, 17, 18] (Fig. 1).

Here in previous three to four days, demand data have been considered as the input time series. Then demand data were tested to check the degree of variation in the statistical data. Finally, autocorrelation factor (ACF) and partial autocorrelation factor (PACF) have been determined. This, in turn, helped to select the value of p and q , white noise for residuals was tested again.

The ARIMA model, which is best suited, would be used to predict the demand for a particular day. AIC and BIC are the general approach for the model selection.

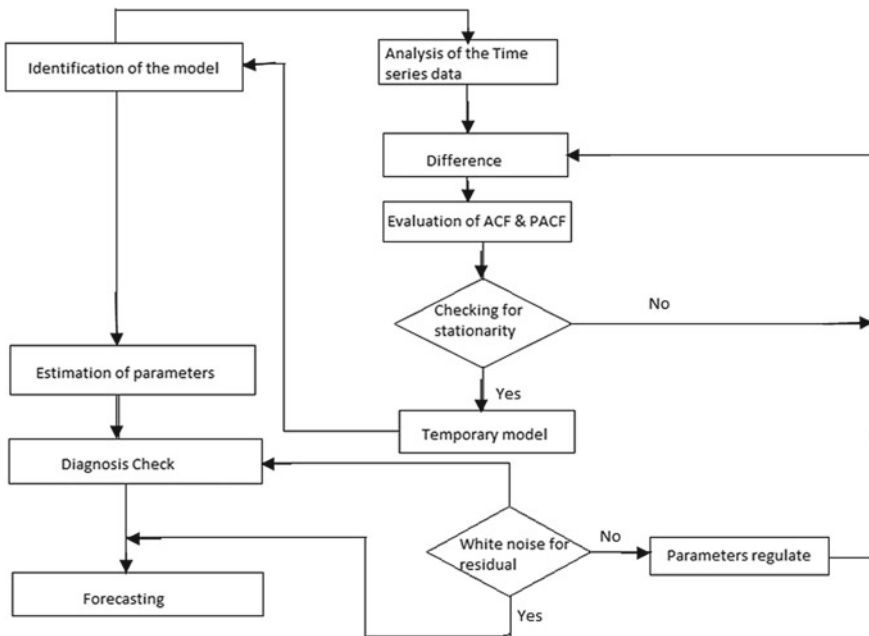


Fig. 1 Flow chart to establish ARIMA model

3.2 ARIMA with Exogenous Variables (ARIMAX)

Different approaches of ARIMAX modelling exist. Conversion of a regression model with autoregression integrated moving average error to a regression integrated moving average with exogenous variables model has been discussed here.

The expression for regression model with ARIMA errors is mentioned in Eq. (1),

$$Y_t = C + \beta X_t + u_t \quad (1)$$

where

Y_t = time series of response variables,

X_t = time series of predictor variables,

C = regression model intercept,

B = regression coefficients,

u_t = time series of independent disturbance.

$$a(L)A(L)(1-L)^D(1-L^S)u_t = b(L)B(L)\varepsilon_t \quad (2)$$

where

ε_t is the innovation series

$$L^j Y_t = Y_{t-j}$$

$$a(L) = 1 - a_1L - \dots - a_pL^p$$

$$A(L) = 1 - A_1L - \dots - A_{ps}L^p$$

$(1-L)^D$, the degree D , non-seasonal integration polynomial

$(1-L^S)$, the degree s , seasonal integration polynomial

$b(L) = 1 + b_1L + \dots + b_qL^q$, the degree q non-seasonal moving average polynomial

$B(L) = 1 + B_1L + \dots + B_{qs}L^q$, the degree qs seasonal moving average polynomial.

Based on the two mentioned methods, ARIMAX model can be designed to achieve an judicious and efficient load forecasting tool for demand side management.

3.3 Artificial Neural Network (ANN)

The importance of back propagation, which was first introduced in 1970s, was fully acknowledged when it was proposed in a paper by Rumelhart et al. [18]. Several

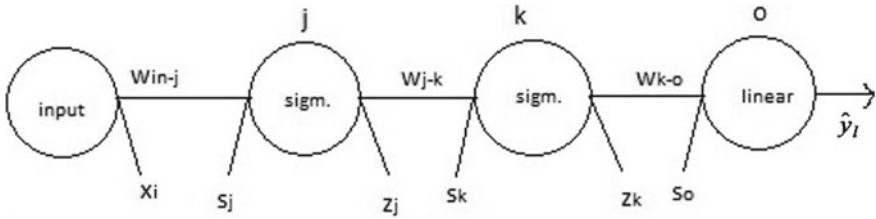


Fig. 2 Network

neural networks have been described in this paper, where back propagation works faster than earlier approaches.

Updating the gradient of each weight with respect to the output is the aim of training the neural network (Fig. 2):

$$w_{ij} = w_{ij} - \rho \frac{\partial E}{\partial w_{ij}} \tag{3}$$

The variables of the shown network are as follows:

$$s_j = w_{1..x_i} \tag{4}$$

$$z_j = \sigma(in_j) = \sigma(w_{1..x_i}) \tag{5}$$

$$s_k = w_{2..z_i} \tag{6}$$

$$z_k = \sigma(in_k) = \sigma(w_{2..z_i}) \tag{7}$$

$$s_o = w_{3..z_k} \tag{8}$$

$$\hat{y}_i = in_o = w_{3..z_k} \tag{9}$$

$$E = \frac{1}{2}(\hat{y}_i - y_i)^2 = \frac{1}{2}(w_{3..z_k} - y_i)^2 \tag{10}$$

The weight updates for the different layers can be found as follows:

$$\frac{\partial E}{\partial w_{ko}} = \frac{\partial}{\partial w_{ko}} \frac{1}{2}(\hat{y}_i - y_i)^2 = \frac{\partial}{\partial w_{ko}} \frac{1}{2}(w_{ko} \cdot z_k - y_i)^2 = (\hat{y}_i - y_i)(z_k) \tag{11}$$

Similarly, the updates of the other layers can be written as follows:

$$\frac{\partial E}{\partial w_{jk}} = (\hat{y}_i - y_i)(w_{ko})(\sigma(s_k)(1 - \sigma(s_k)))(z_i) \tag{12}$$

$$\frac{\partial E}{\partial w_{ij}} = (\hat{y}_i - y_i)(w_{ko})(\sigma(s_k)(1 - \sigma(s_k)))(w_{jk})(\sigma(s_j)(1 - \sigma(s_j)))(x_i) \tag{13}$$

So, finally we can conclude that

$$\Delta w_{ij} = -\rho [(\hat{y}_i - y_i)(w_{ko})(\sigma(s_k)(1 - \sigma(s_k)))(w_{jk})(\sigma(s_j)(1 - \sigma(s_j)))(x_i)] \tag{14}$$

$$\Delta w_{jk} = -\rho [(\hat{y}_i - y_i)(w_{ko})(\sigma(s_k)(1 - \sigma(s_k)))(z_i)] \tag{15}$$

$$\Delta w_{ko} = -\rho [(\hat{y}_i - y_i)(z_k)] \tag{16}$$

The hourwise load data of 2017 and average temperature of the state of West Bengal have been used as learning data for the neural network. A MATLAB code has been generated utilizing the ANN toolbox in MATLAB to forecast the week ahead load demand. The multilayer feed forward network architecture has three types of layers consisting of two hidden layers and one output layer (Fig. 3).

The neuron numbers were changed from 5 to 21 to obtain the model which provides maximum efficiency. The best output was provided by using 12 units in the first layer that is hidden and using 21 units for the second layer that is hidden. The output layer of the neural network which had 24 neurons provided the forecasted electrical load values of each hour for the target day. The historical information of

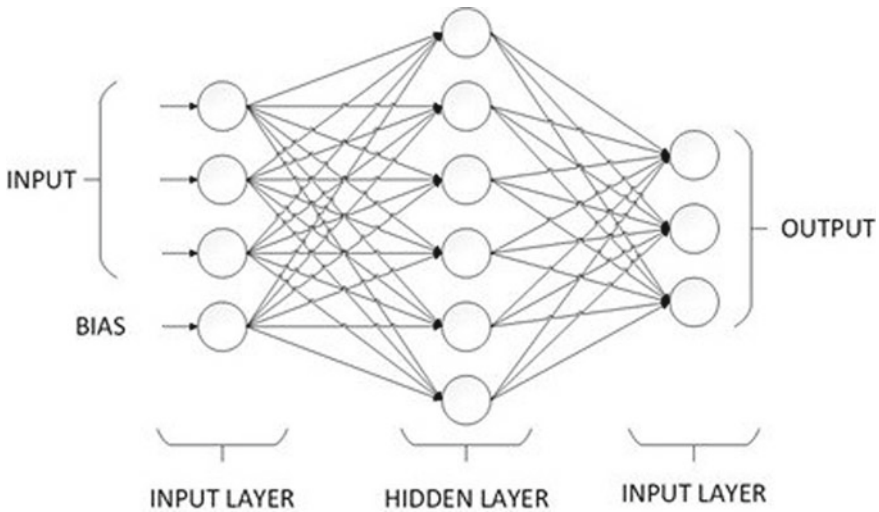


Fig. 3 Architecture of ANN

Table 1 Input and output structures of ANN

Input neurons	Details of input
1, 2, 3, 4, 5, 6, 7, ..., 24	L (day-7, hour); hour = 1, 2, ..., 24
Output neurons	Details of output
1, 2, 3, 4, 5, 6, 7, ..., 24	L (day, hour); h = 1, 2, ..., 24

the electrical load data and other relevant information about the weather or the type of day are provided to the model through the input neurons [19–22].

The details of the input and output structure of ANN are framed below (Table 1).

The scaling of the inputs is done using the ‘prestd’ function so that the input neurons do not get saturated. For getting the actual values, the outputs obtained from the output neurons are reverse scaled to the actual range by using the function ‘poststd’ of MATLAB. The training of the network was done by back propagation algorithm for which the hidden layers and output layers were assigned the log sigmoid and linear activation functions, respectively.

3.4 Long Short-Term Memory (LSTM) Network

Basic architecture of LSTM network is shown in Fig. 4. The input gate (*i*), forget gate (*f*), cell candidate (*g*) and output gate (*o*) control the cell state and hidden state of the layer.

The above components are described at time step *t* by the following formulae:

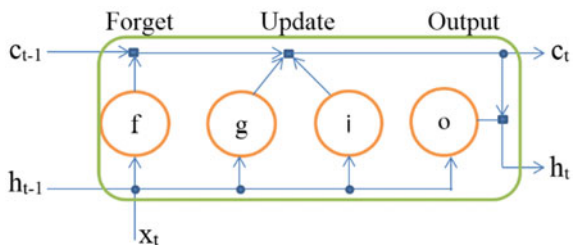
$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \tag{17}$$

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \tag{18}$$

$$g_t = \sigma_c(W_g x_t + U_g h_{t-1} + b_g) \tag{19}$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \tag{20}$$

Fig. 4 LSTM network architecture



where W , U and b denote input weight, recurrent weight and bias, respectively. σ_c and σ_g denote the state activation function and gate activation function, respectively. Here hyperbolic tangent function (\tanh) and sigmoid function $\sigma(x) = (1 + e^{-x})^{-1}$ have been used to compute σ_g and σ_c , respectively.

4 Result and Analysis

To implement the approach that has been proposed in this paper, the demand data of West Bengal, obtained from the website of regional centre of load dispatch, were utilized as the input time series. The hourly and peak load data from January 2017 to December 2017 have been used to check all three models.

Here, both a day and week ahead, load forecasting has been performed. The date for which the week ahead forecasting was performed is Thursday, 31 August, 2017. To achieve an accurate prediction, here past demand data of Thursdays (i.e. 3, 10, 17 and 24 August 2017 and so on) have been used as historical data for ARIMA, ARIMAX or test series data for ANN and LSTM. The day ahead forecasting has been done for 22 September 2017, Friday. To have as much as accurate forecasts, here the daily load demand data of previous days have been used as historical data for ARIMA, ARIMAX, ANN and LSTM. The order of the ARIMAX and ANN model has been chosen on the basis of ACF and PACF and the number of hidden neurons, respectively.

From Figs. 5 and 6, it can be concluded that, in case of week ahead load forecasting, ARIMAX performs better than ARIMA for multivariate analysis.

From Figs. 7, 8, 9 and 10, it can be concluded that the prediction errors have been improved by considering the weather effect (temperature) in case of ANN and LSTM, respectively.

Figures 11 and 12 show that ARIMAX gives more accurate result than ARIMA for a day ahead load forecasting. In Fig. 13, the prediction error has been improved (Figs. 14, 15 and 16; Table 2).

The forecasting error has been tabulated here.

For better assessment, we have also calculated mean absolute percentage error (MAPE).

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{\text{Actual} - \text{Forecasted}}{\text{Actual}} \right| \quad (28)$$

where n = number of forecasted points

Mean absolute error (MAE) and MAPE are calculated to measure the amount by which two continuous variables which differs from each other.

All codings of the ARIMA, ANN, LSTM and ARIMAX have been written in MATLAB environment. The reduction of error has been done for all the methods by multivariate analysis over univariate analysis. Further reduction of error can be

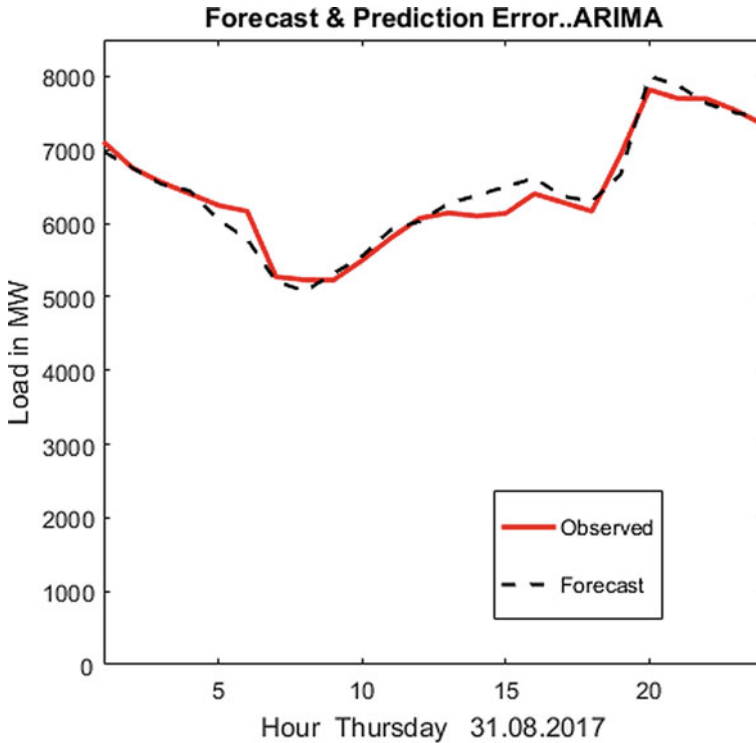


Fig. 5 Actual and forecasted demand profile for 31 August (Thursday) 2017 using ARIMA

done by incorporating several weather conditions into the prediction model. A very short-term, i.e. forecasting of load data for a few minutes ahead of schedule is also possible using the above methods which in turn may help the load forecasting in real time.

5 Conclusion

Electrical load forecasting is utmost important for demand side management. The efficiency of an energy management system depends on load forecasting and thereby renewable energy integration also. To maintain an uninterrupted, good quality power supply during peak hours, it is essential to cater the load from conventional energy sources to renewable energy sources on the basis of the characteristics of the load and forecasted demand profile. There is a high requirement of an efficient forecasting tool. In this article, we have introduced three models of ARIMAX, ANN and LSTM as univariate and multivariate methods to forecast electrical load data and find out the comparative results of these three methods. We can see that all these methods have the

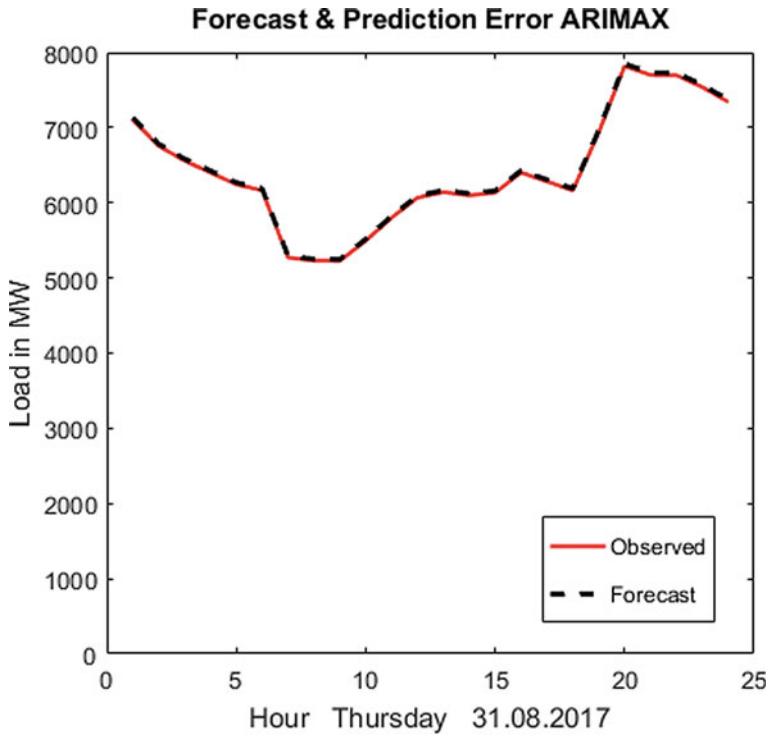


Fig. 6 Actual and forecasted demand profile for 31 August (Thursday) 2017 using ARIMAX

capability to outperform the other depending on the conditions under which they are applied. Our research reveals that ARIMAX, ANN and LSTM achieve remarkable accuracy in peak load as well as daily demand forecasting. The main advantage of using a neural network model as a tool for electrical load forecasting is that the internal architecture of the model need not be modified during the process. Different ARIMAX models are required for variations in the input, as for ARIMAX, the order of prediction model depends on ACF and PACF of past and present data. To design an accurate model for future load forecasting using ARIMAX is a tedious job. But this is more dependable pertaining to weather variables in case of daily demand profile prediction. Here in this article, from Figs. 5 and 12, it is clear that ARIMAX is also solving the purpose of peak load forecasting with better accuracy. Therefore, it can be concluded that in future ARIMAX may prove itself as a highly efficient tool to serve peak load forecasting and real-time forecasting.

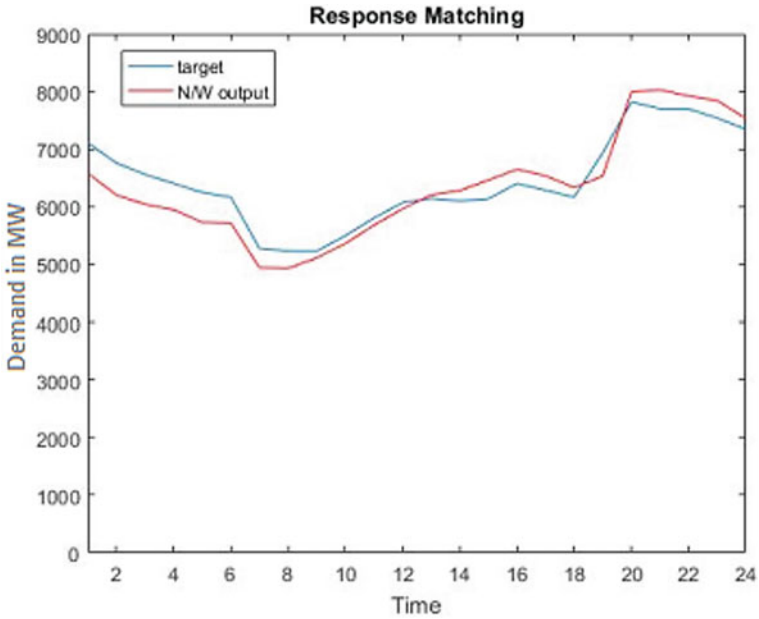


Fig. 7 Actual and forecasted demand profile for 31 August (Thursday) 2017 using univariate ANN

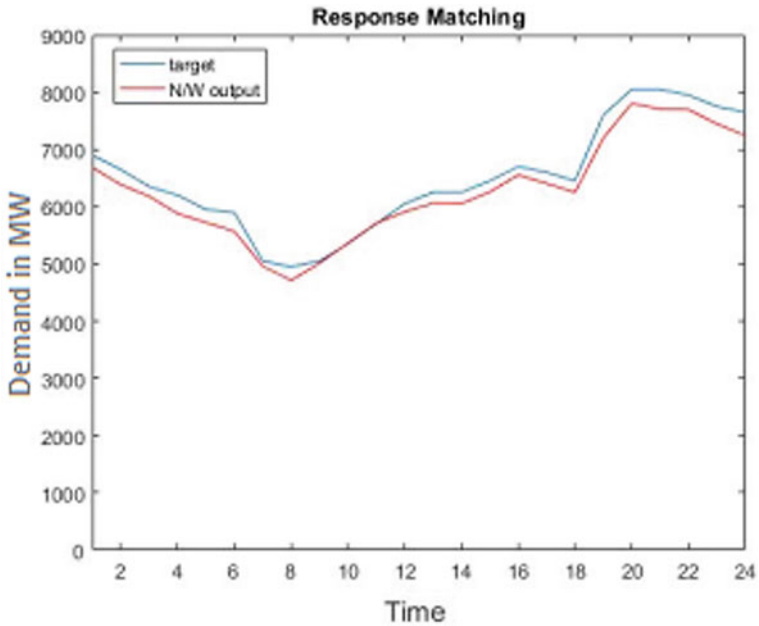


Fig. 8 Actual and forecasted demand profile for 31 August (Thursday) 2017 using multivariate ANN

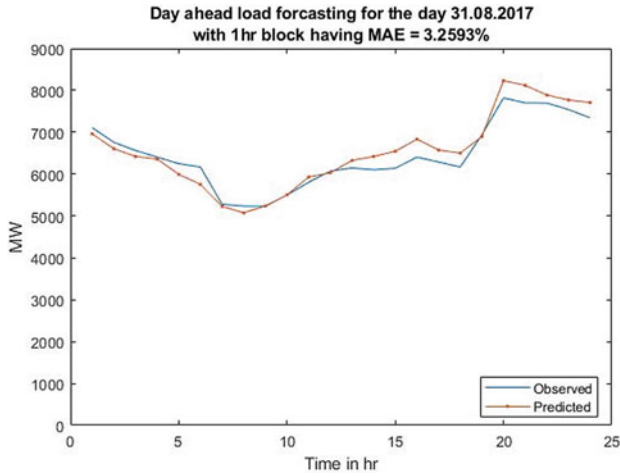


Fig. 9 Actual and forecasted demand profile for 31 August (Thursday) 2017 using univariate LSTM network

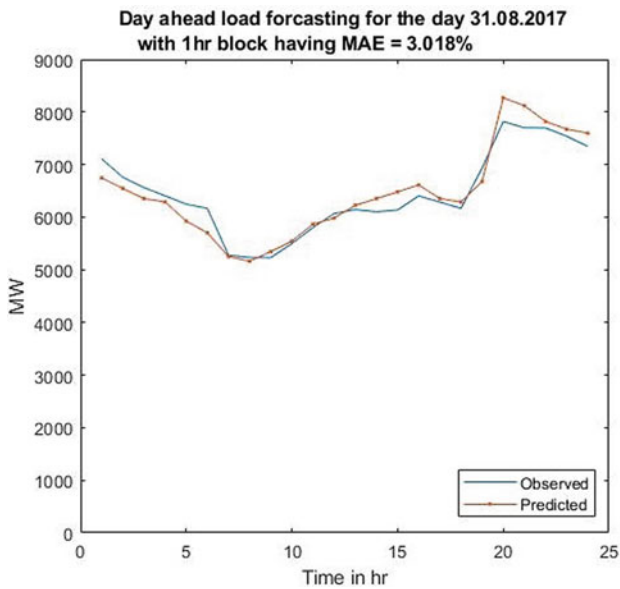


Fig. 10 Actual and forecasted demand profile for 31 August (Thursday) 2017 using multivariate LSTM network

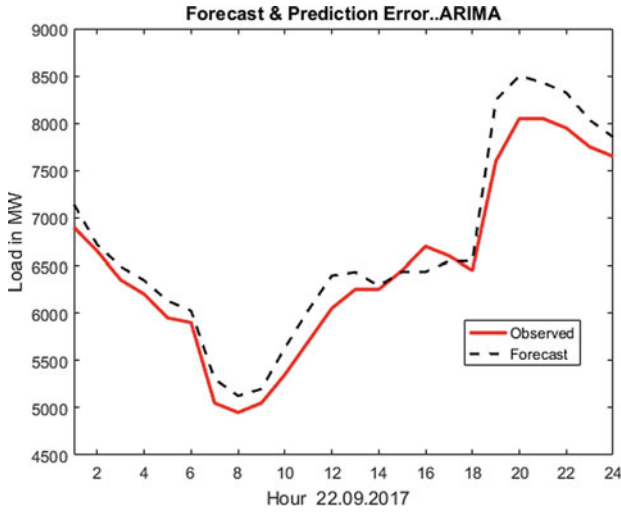


Fig. 11 Actual and forecasted demand profile for 22 September 2017 using ARIMA

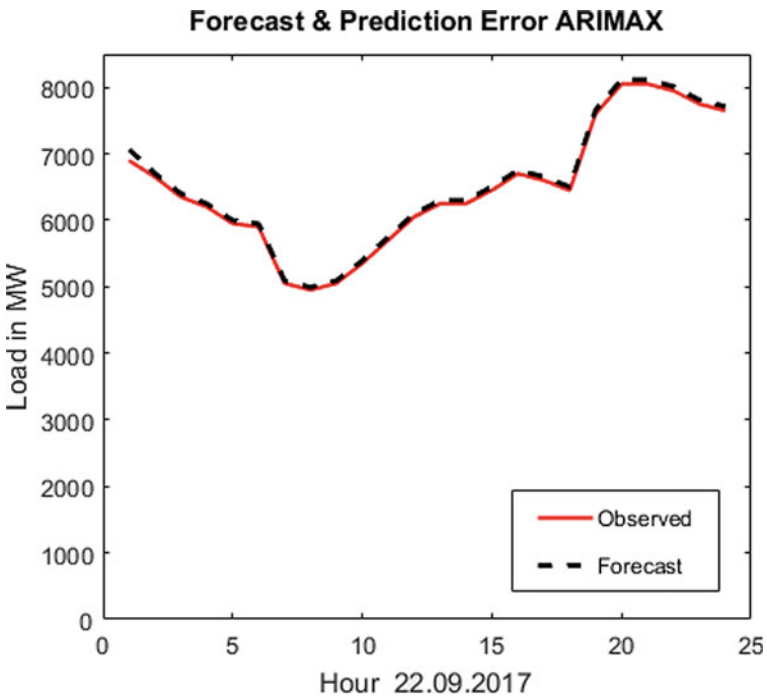


Fig. 12 Actual and forecasted demand profile for 22 September 2017 using ARIMAX

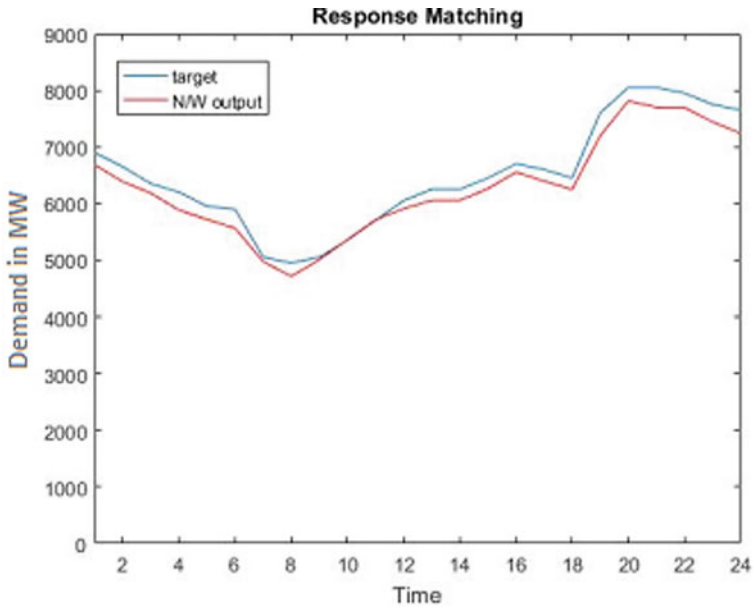


Fig. 13 Actual and forecasted demand profile for 22 September 2017 using univariate ANN

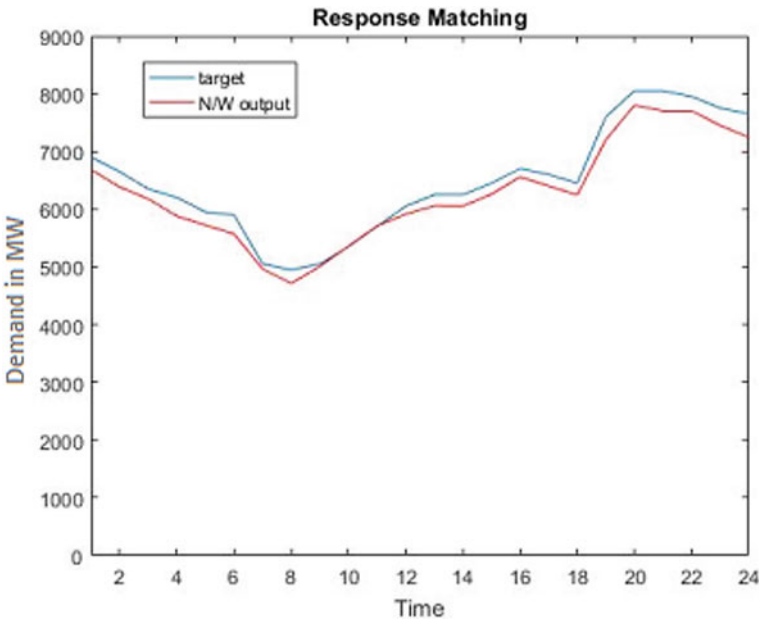


Fig. 14 Actual and forecasted demand profile for 22 September 2017 using multivariate ANN

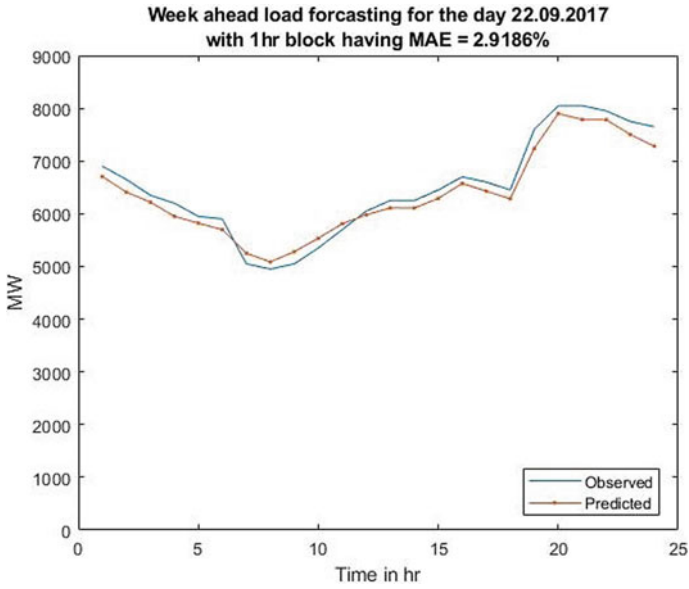


Fig. 15 Actual and forecasted demand profile for 22 September 2017 using univariate LSTM network

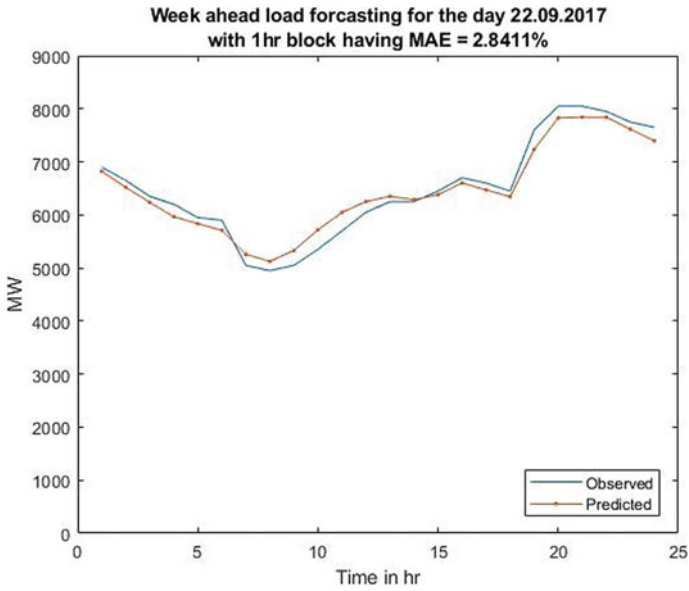


Fig. 16 Actual and forecasted demand profile for 22 September 2017 using multivariate LSTM network

Table 2 Obtained forecasted results

Date and day	Day type	Type of forecasting	Univariate method (without considering temperature)			Multivariate method (considering temperature)		
			Prediction error (MAPE) using ANN (in %)	Prediction error (MAPE) using ARIMA (in %)	Prediction error (MAPE) using LSTM (in %)	Prediction error (MAPE) using ANN (in %)	Prediction error (MAPE) using ARIMA (in %)	Prediction error (MAPE) using LSTM (in %)
Thursday 31.08.2017	Weekdays	Week ahead	4.6934	2.1778	3.2593	3.3299	0.4008	3.0186
Friday 22.09.2017	Weekdays	Week ahead	3.3461	2.9668	2.9186	3.3253	1.8973	2.8411

Acknowledgements The authors are thankful to Eastern Region Load Dispatch Centre (ERLDC) for providing daily demand profile. Jadavpur University and its digital library have played the key role to enhance the research quality. Therefore, the authors would also like to thank Jadavpur University for providing the best study and research environment.

References

1. A.-K. Srivastava, A.-S. Pandey, D. Singh, Short-term load forecasting methods: a review, in *International Conference on Emerging Trends in Electrical Electronics & Sustainable Energy Systems (ICETEESSES)*, Sultanpur, pp. 130–138 (2016)
2. N. Amjady, Short-term hourly load forecasting using time-series modeling with peak load estimation capability. *IEEE Trans. Power Syst.* **16**(3), 498–505 (2001)
3. H. Daneshi, *Real Time Load Forecast in Power Systems*. DRPT2008, Nanjing, China (2008)
4. D. Ross, T. Athay, S. Kim, G. Ackerman, R. Podmore, S. Virmani, *Development and Implementation of Advanced Automatic Generation Control*. U.S. Dept. Energy Rep. EC-77-01-2118(1978)
5. K.-H. Kim, H.-S. Youn, Y.-C. Kang, Short-term load forecasting for special days in anomalous load conditions using neural networks and fuzzy inference method. *IEEE Trans. Power Syst.* **15**, 559–565 (2000)
6. T.-M. Peng, N.-F. Hubele, G.-G. Karady, An adaptive neural network approach to one-week ahead load forecasting. *IEEE Trans. Power Syst.* **8**, 1195–1203 (1993)
7. T. Senjyu, P. Mandal, K. Uezato, T. Funabashi, Next day load curve forecasting using recurrent neural network structure, in *Generation, Transmission and Distribution, IEEE Proceedings*, vol. 151, pp. 388–394 (2004)
8. J.-W. Taylor, R. Buizza, Neural network load forecasting with weather ensemble predictions. *IEEE Trans. Power Syst.* **17**, 626–632 (2002)
9. J. Schachter, P. Mancarella, A short-term load forecasting model for demand response applications. 978-1-4799-6095-8/14/\$31.00 © 2014 IEEE
10. A. Azadeh, S.-F. Ghadrei, B.-P. Nokhandan, One day ahead load forecasting for electricity market of Iran by ANN, in *2009 International Conference on Power Engineering, Energy and Electrical Drives*, Lisbon, pp. 670–674 (2009)
11. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
12. G. Juberias, R. Yunta, J.-G. Moreno, A new ARIMA model for hourly load forecasting, in *Transmission and Distribution Conference, 1999 IEEE* (2002)
13. R. Kelly, R. Modarres, A. Sarhadi, Snow water equivalent time series forecasting in Ontario, Canada, in link to large atmospheric circulations. *Hydrol. Process.* **28**(16), 4640–4653 (2014)
14. Y. Belkhier, A. Achour, N. Ullah, R.N. Shaw. Modified passivity-based current controller design of permanent magnet synchronous generator for wind conversion system. *Int. J. Model. Simul.* (2020). <https://doi.org/10.1080/02286203.2020.1858226>
15. Y. Belkhier, A. Achour, R.N. Shaw, Fuzzy passivity-based voltage controller strategy of grid-connected PMSG-based wind renewable energy system, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, 2020, pp. 210–214. <https://doi.org/10.1109/iccca49541.2020.9250838>
16. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh (eds) *Innovations in Electrical and Electronic Engineering*. Lecture Notes in Electrical Engineering, vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58

17. K.-B. Sahay, M.-M. Tripathi, Day ahead hourly load forecast of PJM electricity market and iso New England market by using artificial neural network, in *ISGT 2014*, Washington, DC (2014)
18. R.N. Shaw, P. Walde, A. Ghosh, A new model to enhance the power and performances of 4×4 PV arrays with puzzle shade dispersion. *Int. J. Innov. Technol. Explor. Eng.* **8**(12) (2019)
19. S. Muzaffar, A. Afshari, Short-term load forecasts using LSTM networks. in presented at *10th International Conference on Applied Energy*, Hong Kong (2018)
20. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON) held at Deenbandhu Chhotu Ram University of Science and Technology, SONEPAT*. India on Feb 28–March 1, 2020
21. B. Kwon, R. Park, K. Song, Short term load forecasting based on deep neural networks using LSTM layer. *J. Electr. Eng. Technol* **15**, 1501–1509 (2020)
22. Ermatita et al.: Peak load forecasting based on long short term memory, in presented at *International Conference on Informatics, Multimedia, Cyber and Information System*, Jakarta, Indonesia (2019)

Comparative Performance of Different PV Array Topologies Under Partial Shading Condition



Karni P. Palawat, Vinod K. Yadav, Ramjee L. Meena, and Santosh Ghosh

Abstract The world has resorted to photovoltaic (PV) technologies to alleviate the broad spectrum of issues posed by fossil fuel-based energy generation, and presently global cumulative PV capacity has crossed over 650 GW. However, PV technology has some inherent challenges which need to be addressed for effective and efficient utilization. Partial shading condition (PSC) is one of those, which does not only diminish the peak power output but also causes hot spotting and irreversible damage of PV cells in some cases. In the present work, the comparative performance of different PV topologies is studied, under various PSCs simulated in MATLAB[®]/Simulink platform.

Keywords Partial shading condition · Photovoltaics · PV array topology · Modeling and simulation · MATLAB[®]

1 Introduction

Presently, the policymakers, researchers, and investors are working toward bringing a paradigm shift in energy generation globally, from conventional to renewable technologies to mitigate global warming and climate change [1, 2]. And due to this joint effort of all the stakeholders, a technological revolution has taken place over the last two decades, which has made PV energy more economical than all other energy conversion technologies ever known to mankind [3]. However, there are some inherent challenges associated with PV technologies, such as PSCs and aerosol

K. P. Palawat · V. K. Yadav · R. L. Meena
Department of Electrical Engineering, Delhi Technological University, Delhi 110042, India
e-mail: vinodkumar@dtu.ac.in

R. L. Meena
e-mail: R.l.meena@dce.edu

S. Ghosh (✉)
Corporate R&D Department, Kirloskar Brothers Limited, Pune 411045, India
e-mail: Santosh.ghosh@kbl.co.in

deposition, which are the barrier against its utilization to its full potential [4]. PSC impacts the PV module performance in multiple ways. In addition to global maximum power point (MPP) (GMPP), a number of local MPP (LMPP) are also produced in the performance characteristics [5], which makes maximum power point tracking (MPPT) extremely challenging. Secondly, under PSC, the affected cells act as a load instead of a photocurrent generator and are reverse biased; hence, the temperature of those cells increases [6]. Due to this phenomenon, the efficiency and power output drop further, and the reliability of the PV modules is impacted as well. Due to reverse breakdown of the affected cells, secondary breakdown or thermal breakdown occurs, leading to a rise in temperature to as higher as 400 °C [7] in some cases, which eventually results in irreversible damage of the PV modules [8]. The PSCs may be created by dust accumulations, bird droppings, shadows of nearby structures, fallen leaves, etc. [9]. Several solutions are proposed to mitigate the impact of PSC on performance and reliability by the researchers in the past, such as innovative hotspot mitigating circuit [7], module-level DC–DC converters [10], and inverters [11], differential converters, and innovative array topologies [12]. In the present work, the comparative performance of three such array topologies, which are series–parallel (SP), bridge link (BL), and total cross-tied (TCT), are studied under varying shading patterns and intensities, in MATLAB®/Simulink platform.

2 Modeling of PV Array

The PV arrays are constructed by interconnecting strings of PV modules in different ways, based upon the ratings of the modules and the plant capacity. The PV strings are, basically, a number of PV modules connected in series and the length of strings (i.e., the number of modules connected in series) depends upon the design voltage rating of the plant, which is generally either 1000 or 1500 V. Further, the number of such PV strings is connected in parallel configuration based upon the design plant capacity in terms of power output, which in turn depends upon the land availability. Instead of providing only two busses at both the ends of the strings (as in SP configuration), multiple cross-ties are provided to form different PV array configuration with higher interconnection redundancy (see Fig. 1). Though these additional cross-ties increase cost, they provide redundant paths for the generated current in the scenario of PSC or module failures [13]. Hence, the trade-off between the degree of interconnection redundancy and improvement of performance under PSC is important. In the present work, comparative performance of three array configurations is studied, which have a different degree of interconnection redundancy, under different shading patterns and shading intensities.

As discussed in the paragraph above, the basic constituent element of an array is a PV module, and each PV module is manufactured by connecting a number of PV cells in series, depending upon the rated power output of the module. Hence, the complete PV arrays of different topologies of interest can be modeled by interconnecting the

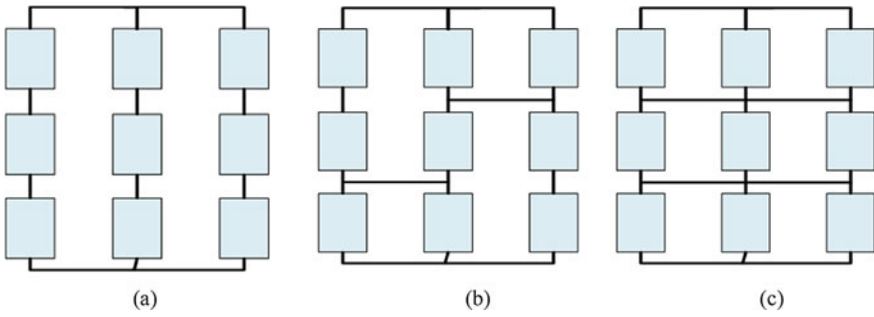


Fig. 1 PV array configurations studied **a** series–parallel, **b** honeycomb, and **c** total cross-tied

model of PV cells which is its basic constituent element. The fundamental equations of the PV cell model are briefly discussed in the following section.

2.1 Modeling of PV Cells

The current versus voltage ($I-V$) and power versus voltage ($P-V$) characteristics of PV cells may be represented by a single diode equivalent circuit model (refer to Fig. 2). By solving the equivalent circuit model shown in Fig. 2, the output of the

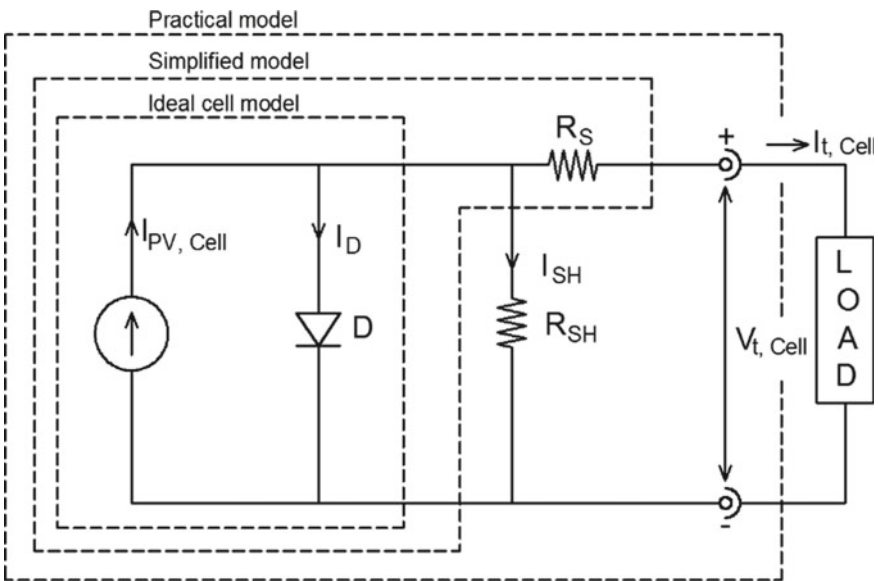


Fig. 2 Equivalent circuit model (single diode) of PV cell

PV cell in terms of current can be obtained as follows:

$$I_{t,\text{cell}} = I_{\text{PV,cell}} - I_D - I_{\text{SH}} \tag{1}$$

where

$$I_D = I_r \left[\exp \left(\frac{V_{t,\text{cell}} + R_s I_{t,\text{cell}}}{V_{t,\text{cell}} \propto} \right) - 1 \right] \tag{2}$$

and current through shunt resistance,

$$I_{\text{SH}} = \frac{V_{t,\text{cell}} + R_s I_{t,\text{cell}}}{R_{\text{sh}}} \tag{3}$$

Substituting (2) and (3) in (1), the nonlinear relation between the PV cell output voltage and current can be mathematically modeled as follows:

$$I_{t,\text{cell}} = I_{\text{PV,cell}} - I_r \left[\exp \left(\frac{V_{t,\text{cell}} + R_s I_{t,\text{cell}}}{V_{t,\text{cell}} \propto} \right) - 1 \right] - \frac{V_{t,\text{cell}} + R_s I_{t,\text{cell}}}{R_{\text{sh}}} \tag{4}$$

$$\text{Here, cell output voltage } V_{t,\text{cell}} = k \frac{T}{q} \tag{5}$$

where k is Boltzmann’s constant ($1.38060503 \times 10^{-33}$ J/K), q is the charge of electrons (1.6×10^{-19} c), T is a cell operating temperature (K), R_s and R_{sh} are series and shunt resistances, respectively, and “ \propto ” stands for diode emission coefficient. The ideal value of \propto is 1. There are different methods to estimate the value of “ \propto ” for optimization of power output, which is discussed in [14] in detail.

The model of the PV module and the different arrays are built by suitably inter-connecting the model of their basic constituent element, i.e. a PV cell, represented by (4). In the present work, model of a 36-cell module was developed for analyzing the performance of an array of 3×3 matrix (see Fig. 1), under simulated PSCs. The specification of the PV module model used for the study is presented in Table 1.

Table 1 Specification of the PV module used for developing the arrays

Parameters	Unit	Value
Maximum power output	W	42.5
Voltage at maximum power point (V_{mpp})	V	18.52
Current at maximum power point (I_{mpp})	A	2.295
Short-circuit current (I_{sc})	A	2.428
Open-circuit voltage (V_{oc})	V	22.65
Number of cells per module	Nos.	36

2.2 Simulated Partial Shading Conditions

The simulated partial shading conditions are shown in Fig. 3. As shown in the said figure, nine different types of shading patterns are simulated, which emulates different shading scenarios of the real world. For example, Fig. 3b, c emulates the direct shading of a pole when the geometrical orientation of the sun's path of transition, pole, and the array are aligned, whereas, Fig. 3f, g emulates the shading of a pole when the orientation of the array is at 45° with respect to the path of the transition of the sun. Figure 3h, i emulates random shading due to fallen leaves or uneven deposition of aerosols. Figure 3d, e emulates the shading of a low height structure, such as the boundary wall and parapet of a roof (for a rooftop installation). The result of the simulation is presented in ensuing Sect. 3.

2.3 Result and Discussion

The $P-V$ and $I-V$ characteristics obtained from the analysis of all the three array topologies (see Fig. 1) are presented in Figs. 4, 5 and 6, which provide deeper insights regarding the behavior of different array topologies under nine different partial shading conditions. It may be observed that multiple peaks were generated in the $I-V$ and $P-V$ characteristics in all the array topologies under PSC, and the global peak value reduced under all the PSCs except under column shading pattern. It is interesting to note that the amplitude of the global peak reduced in varying degrees in the case of different array topologies, under PSCs. In three out of eight simulated PSCs, the maximum power output (i.e., global peaks) is observed to be equal in all the three array topologies, which are column shading (even) and row shading (both even and uneven). For the rest of the five simulated PSCs, TCT topology yielded maximum power output (see Fig. 7), followed by BL and SP topologies, in order, except for diagonal (both even and uneven) shading pattern, in case of which the power loss in SP and BL topologies was equal. It is also fascinating to note that multiple peaks (GMPP and LMPP) were generated by SP and BL topologies under all the shading conditions except column shading; however, TCT configuration exhibited a single peak point in case of diagonal shading pattern as well, apart from column shading (Table 2).

3 Conclusion

In the present work, the comparative performance of three array topologies is analyzed, under eight different shading conditions that emulate different real-world scenarios of partial shading. It is found from the result of the simulation that the average peak power output under all the simulated shading scenarios increased with

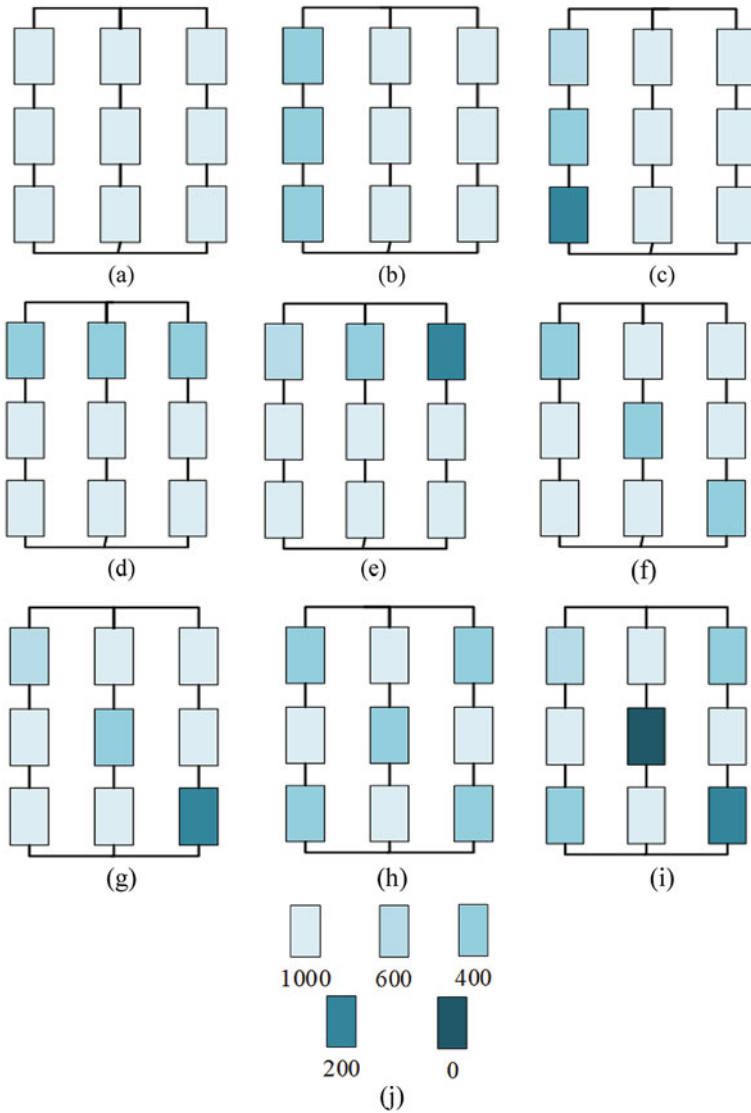


Fig. 3 Simulated shading patterns and intensity **a** non-shaded (NS), **b** column shading (CS), **c** uneven column shading (UCS), **d** row shading (RS), **e** uneven row shading (URS), **f** diagonal shading (DS), **g** uneven diagonal shading (UDS), **h** random shading (RanS), **i** uneven random shading (URanS), and **j** legends of shading intensity. Different shading intensity was simulated by varying the irradiance. The figure represents the simulated shading patterns of different intensities for SP array configuration only as a sample case [15–18]. All the array topologies (see Fig. 1) were subjected to the same set of shading patterns and intensities.

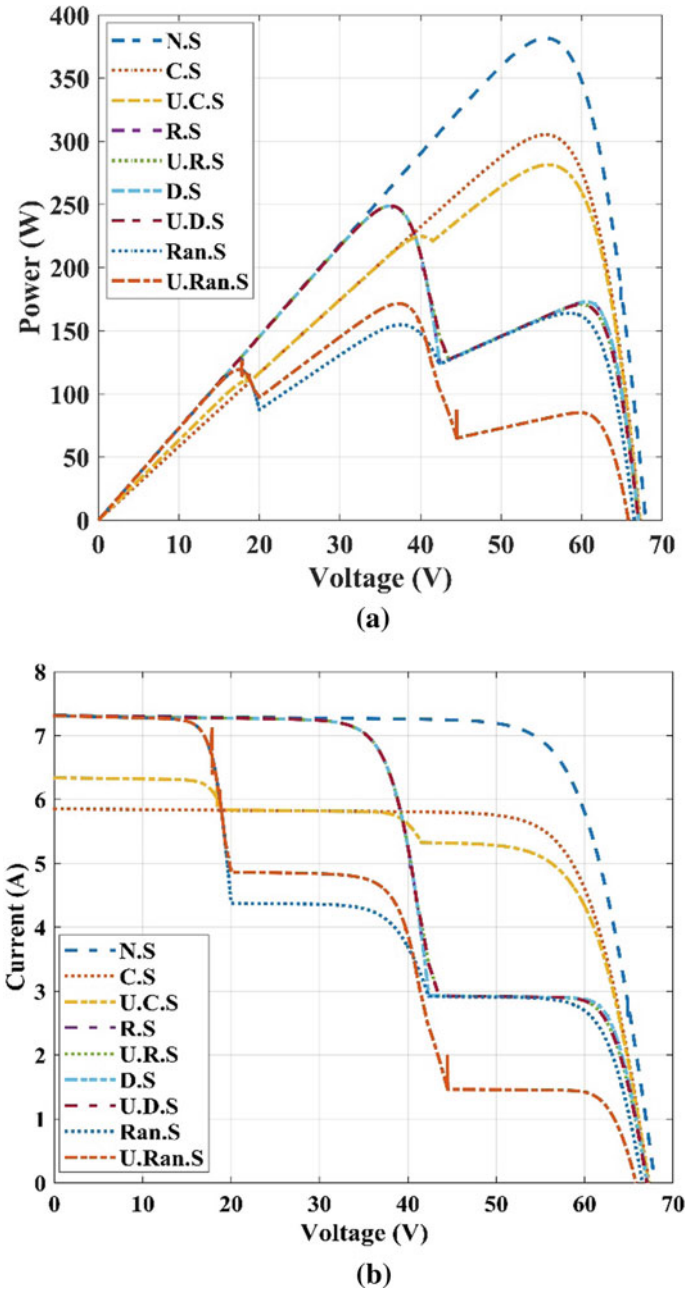
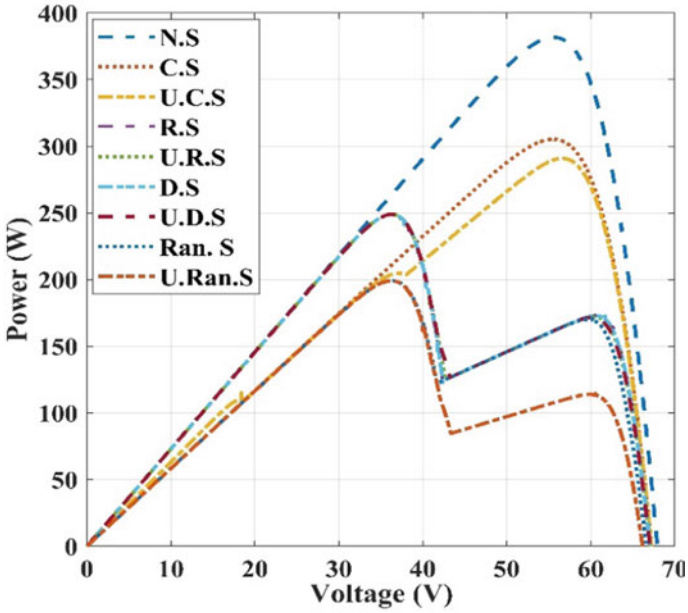
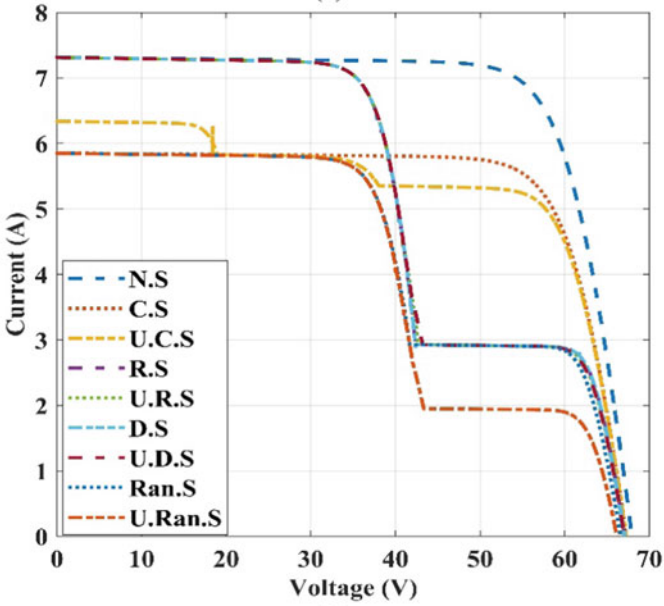


Fig. 4 Characteristics of SP array topology under simulated PSCs **a** $P-V$ and **b** $I-V$



(a)



(b)

Fig. 5 Characteristics of BL array topology under simulated PSCs a $P-V$ and b $I-V$

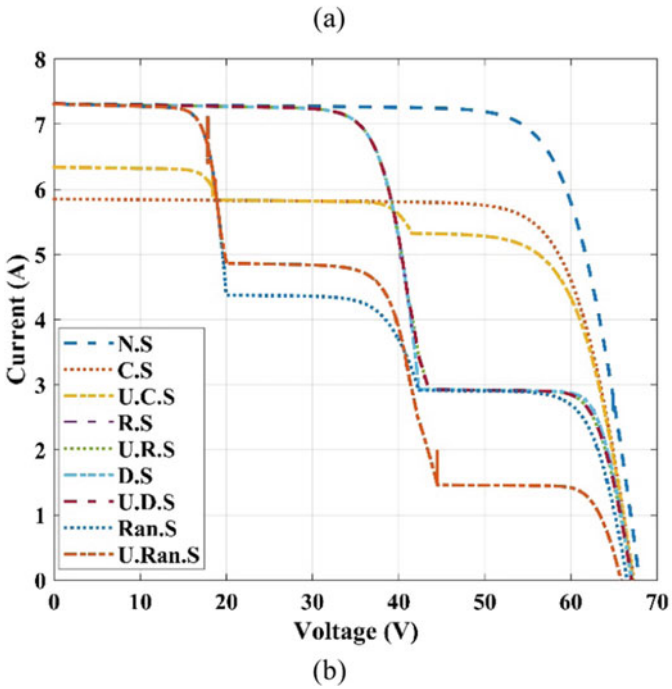
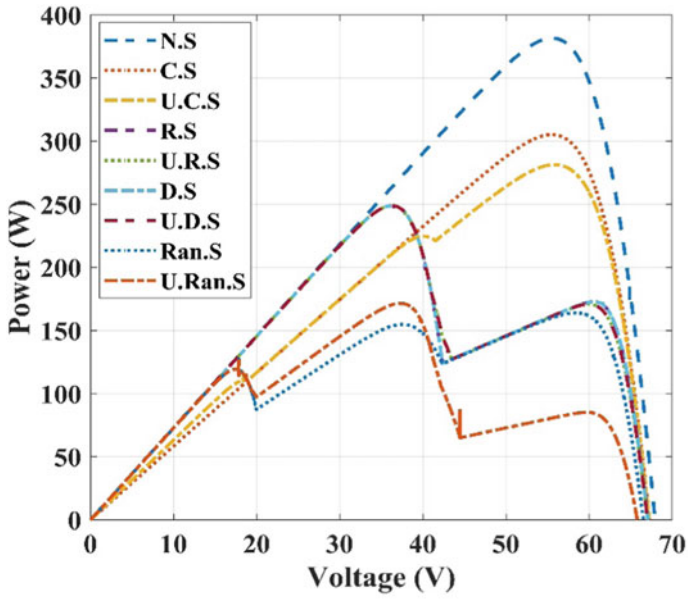


Fig. 6 Characteristics of TCT array topology under simulated PSCs a $P-V$ and b $I-V$

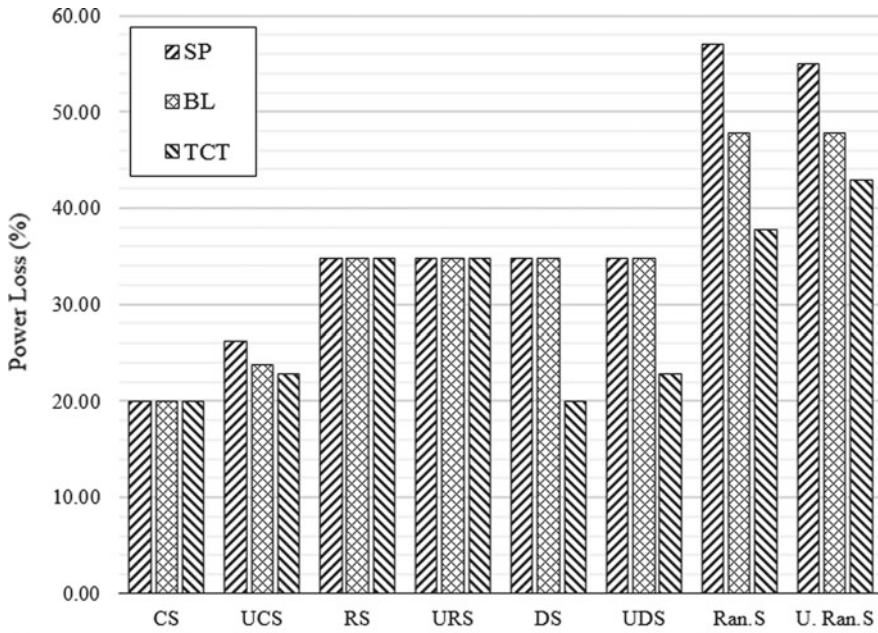


Fig. 7 Comparative power loss in different array topologies under all the simulated PSCs

the increase in interconnection redundancy. The average peak power output of TCT, BL, and SP configuration is found to be 112.5 W, 132.8 W, and 141.8 W, respectively, under eight PSCs considered in the present work. Hence, cost–benefit analysis is recommended for each site during the design stage, considering the increase in cost due to the increase in the number of cross-ties and the reduction in power loss under partial shading conditions, because of those.

Table 2 Parameters of the PV module used for developing the arrays

Topology	PSCs	GMPP			LMPP		
		V_{GMPP} (V)	I_{GMPP} (A)	P_{GMPP} (W)	V_{LMPP} (V)	I_{LMPP} (A)	P_{LMPP} (W)
Series-parallel (SP)	NS	55.78	6.84	381.6	-	-	-
	CS	55.44	5.50	305.2	-	-	-
	UCS	56.01	5.02	281.4	40.9	5.49	224.7
	RS	36.28	6.86	248.9	60.58	2.85	172.9
	URS	36.28	6.86	248.9	59.94	2.85	170.9
	DS	36.28	6.86	248.9	60.58	2.85	172.9
	UDS	36.28	6.86	248.9	59.9	2.85	171
	Ran.S	58.30	2.81	164.0	37.74	4.10	154.8
	U.Ran.S	37.31	4.59	171.9	17.44	6.84	119.3
	NS	55.3	6.9	381.6	-	-	-
Bridge link (BL)	CS	55.31	5.51	305.2	-	-	-
	UCS	55.65	5.22	290.8	37.2	5.50	204.8
	RS	36.28	6.86	248.9	60.57	2.85	172.9
	URS	36.28	6.86	248.9	60.8	2.82	172
	DS	36.28	6.86	248.9	60.58	2.85	172.9
	UDS	36.28	6.86	248.9	60.38	2.84	171.6
	Ran.S	36.3	5.48	199.3	59.58	2.85	170.1
	U.Ran.S	36.32	5.48	199.1	60.54	1.88	114.3
	NS	55.76	6.84	381.6	-	-	-
	CS	55.51	5.49	305.2	-	-	-

(continued)

Table 2 (continued)

Topology	PSCs	GMPP			LMPP		
		V_{GMPP} (V)	I_{GMPP} (A)	P_{GMPP} (W)	V_{LMPP} (V)	I_{LMPP} (A)	P_{LMPP} (W)
	UCS	56.73	5.19	294.5	36.57	5.60	204.9
	RS	36.28	6.86	248.9	60.66	2.85	172.9
	URS	36.28	6.86	248.9	60.66	2.85	172.8
	DS	55.51	5.49	305.2	–	–	–
	UDS	56.69	5.19	294.5	36.58	5.60	204.9
	Ran.S	56.63	4.19	237.4	17.01	5.46	92.96
	U.Ran.S	57.2	3.78	217.9	35.95	4.60	165.5

Cells with no data in LMPP columns indicate only one maximum power point is created under that particular shading pattern and intensity, i.e., no LMPP is created

References

1. A.S. Yadav V.K. Yadav V. Mukherjee S. Ghosh, Performance investigation of different bypass diode topology based SDK-PV arrays under partial shading conditions, in ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh, *Innovations in Electrical and Electronic Engineering*. Lecture Notes in Electrical Engineering, vol. 661, (Springer, Singapore, 2021), pp. 261–270. https://doi.org/10.1007/978-981-15-4692-1_20
2. S. Ghosh, V.K. Yadav, V. Mukherjee, P. Yadav, Evaluation of relative impact of aerosols on photovoltaic cells through combined Shannon's entropy and data envelopment analysis (DEA). *Renew. Energy* **105**, 344–353 (2017)
3. S. Ghosh, V.K. Yadav, V. Mukherjee, R. Birajdar, Status check: journey of India's energy sustainability through renewable sources. *Control of Power Energy Syst., IFAC-PapersOnLine* **48**(30), 456–461 (2015)
4. S. Ghosh, V.K. Yadav, V. Mukherjee, Impact of environmental factors on photovoltaic performance and their mitigation strategies—a holistic review. *Renew. Energy Focus* **28**(00), 153–172 (2019). <https://doi.org/10.1016/j.ref.2018.12.005>
5. Y.H. Liu, S.C. Huang, J.W. Huang, W.C. Liang, A particle swarm optimization-based maximum power point tracking algorithm for PV systems operating under partially shaded conditions. *IEEE Trans. Energy Convers.* **27**(4), 1027–1035 (2012). <https://doi.org/10.1109/TEC.2012.2219533>
6. S. Ghosh, V.K. Yadav, V. Mukherjee, Improvement of partial shading resilience of PV array though modified bypass arrangement. *Renew. Energy* **143**, 1079–1093 (2019)
7. S. Ghosh, V.K. Yadav, V. Mukherjee, A novel hot spot mitigation circuit for improved reliability of PV module. *IEEE Trans. Device Mater. Reliab.* **20**(1), 191–198 (2020). <https://doi.org/10.1109/TDMR.2020.2970163>
8. S. Ghosh, V.K. Yadav, V. Mukherjee, Evaluation of cumulative impact of partial shading and aerosols on different PV array topologies through combined Shannon's entropy and DEA. *Energy* **144**, 765–775 (2018)
9. F. Bayrak, G. Ertürk, H.F. Oztop, Effects of partial shading on energy and exergy efficiencies for photovoltaic panels. *J. Clean. Prod.* **164**, 58–69 (2017). <https://doi.org/10.1016/j.jclepro.2017.06.108>
10. G. Adinolfi, G. Graditi, P. Siano, A. Piccolo, Multi objective optimal design of photovoltaic synchronous boost converters assessing efficiency, reliability, and cost savings. *IEEE Trans. Ind. Informat.* **11**(5), 1038–1048 (2015)
11. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
12. R.N. Shaw, P. Walde, A. Ghosh, A new model to enhance the power and performances of 4 × 4 PV arrays with puzzle shade dispersion. *Int. J. Innov. Technol. Explor. Eng.* **8**(12) (2019)
13. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON) held at Deenbandhu Chhotu Ram University of Science and Technology, SONEPAT*. India on Feb 28–March 1 2020
14. D. Saravanan, D. Singh, S. Sinha, R.N. Shaw, K. Prabu, A comprehensive study of well-known maximum power point tracking techniques, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, 2020, pp. 829–837. <https://doi.org/10.1109/gucon48875.2020.9231219>
15. S. Harb, R.S. Balog, Reliability of candidate photovoltaic module-integrated-inverter (pv-mii) topologies—a usage model approach. *IEEE Trans. Power Electron.* **28**(6), 3019–3027 (2013)
16. N. Mishra, A.S. Yadav, R. Pachauri, Y.K. Chauhan, V.K. Yadav, Performance enhancement of PV system using proposed array topologies under various shadow patterns. *Sol. Energy* **157**, 641–656 (2017). <https://doi.org/10.1016/j.solener.2017.08.021>

17. K.P. Palawat, V.K. Yadav, R.L. Meena, S. Ghosh, Experimental investigation of performance of PV array topologies under simulated PSCs, in *Applied Soft Computing and Embedded System Applications in Solar Energy*, CRC Press-Accepted
18. G. Walker, Evaluating MPPT converter topologies using a matlab PV model. *J. Electr. Electron. Eng. Aust.* **21**(1), 49–55 (2001)

Fault Detection in Power Transmission System Using Reverse Biorthogonal Wavelet



Gaurav Kapoor, Vishesh Kumar Mishra, Rabindra Nath Shaw,
and Ankush Ghosh

Abstract A fault detection scheme is presented in this paper based on reverse biorthogonal wavelet (RBW). The usefulness of the RBW-based scheme has been tested and authenticated on power transmission system connected with distributed generation. Simulation effects exhibit that very accurate fault detection using the features extracted from the fault currents is feasible.

Keywords Wavelet transform · Transmission line · Fault detection · Power system

1 Introduction

The main duties of transmission line fault detector are awareness of the fault, categorizing the fault type, and locating the fault. Various techniques have been reported to execute these duties. Faults in TL are detected and classified using DWT and double-channel extreme learning machine [1]. In [2], HIFs in power distribution system are detected and classified using LWT, ANN, and ALO. In [3], fault classification task in distribution network is done using CNN. Further, the techniques based on wavelet correlation modes [4], WT with ChNN [5], and DWT with ANN [6] are also introduced for fault classification in TLs. Augmented global relay parameter and normalized feature value are utilized for fault discrimination in TLs [7]. In [8], a decentralized method to wide-area back-up protection of TLs is proposed. Also, ANN-based method [9] and sequential components-based scheme [10] are proposed for fault identification in TLs.

G. Kapoor (✉)

Department of Electrical Engineering, Modi Institute of Technology, Kota, India
e-mail: gaurav.kapoor019@gmail.com

V. K. Mishra

School of EEC Engineering, Galgotias University, Greater Noida, India

R. N. Shaw · A. Ghosh

School of EAS, The Neotia University, Sarisha, West Bengal, India
e-mail: r.n.s@ieee.org

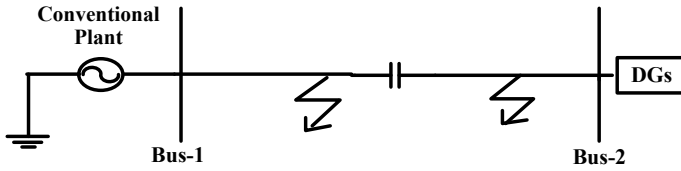


Fig. 1 Presented power system

In this work, the RBW-based method is proposed for fault identification for the power transmission system connected with DGs. An RBW-based method is presented using statistics samples of current signals.

The rest of the paper is prearranged as: In Sect. 2, a depiction of studied system is demonstrated. Section 3 introduces the method of fault identification using RBW. Section 4 presents the effects of various simulations. Finally, the conclusions are presented in Sect. 5.

2 Presented Power System

A power system model is proposed in this paper as demonstrated in Fig. 1. The model (400-kV) is implemented in MATLAB. The general span of line is deliberated as 200-km with frequency is selected as 50-Hz. Bus 1 is fed from conventional plant and Bus 2 is fed from distributed generation (wind generators) at both end points of power system [11–13].

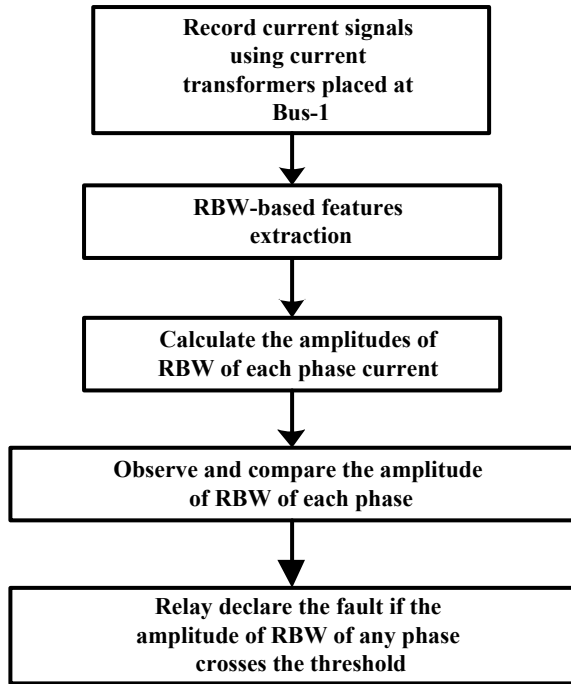
3 The Method of Fault Identification Using RBW

The method of fault identification using RBW is demonstrated in Fig. 2.

4 Simulation Effects

For numerous fault happenings, the model is simulated and the method is tested. The consequences are demonstrated under.

Fig. 2 Method of fault recognition using RBW



4.1 Faultless Condition

The power system current and voltage signals considering healthy operating situation are presented in Fig. 3. The RBW plot is demonstrated in Fig. 4 and the effects are deliberated in Table 1 [14–16].

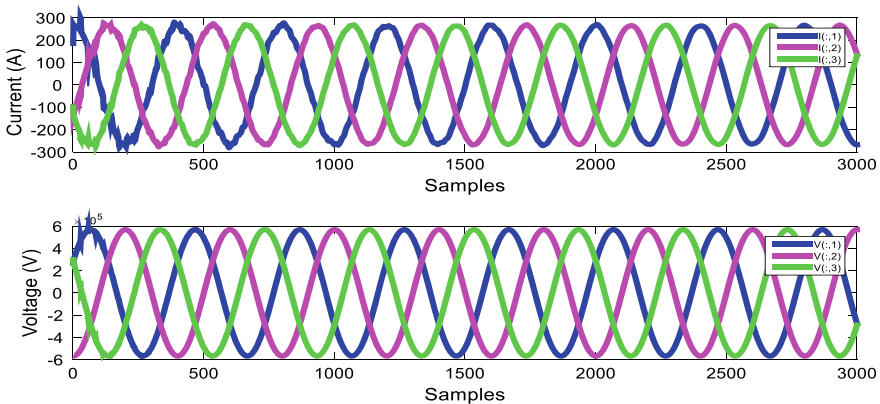


Fig. 3 Power system current and voltage signals

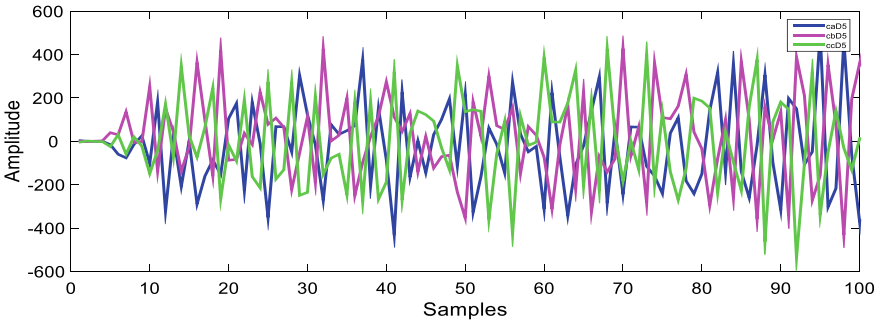


Fig. 4 RBW output of power system current and voltage signals

Table 1 Results deliberating faultless condition

RBW output		
A	B	C
592.7185	470.5970	425.5974

4.2 Evolving Faults

The current waveform during evolving fault (ABG at 0.05 s to BG at 0.15 s) analyzed is demonstrated in Fig. 5 and the obtained RBW output is presented in Fig. 6. The calculated amplitudes of RBW during quite a few faults are reported in Table 2. The presentation of RBW is established moderately satisfactory according to the results reported in Table 2 [17–19].

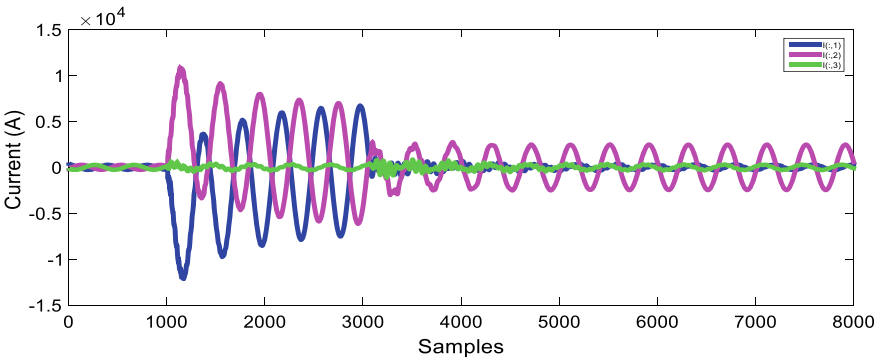


Fig. 5 ABG fault at 0.05 s is evolved to BG fault at 0.15 s

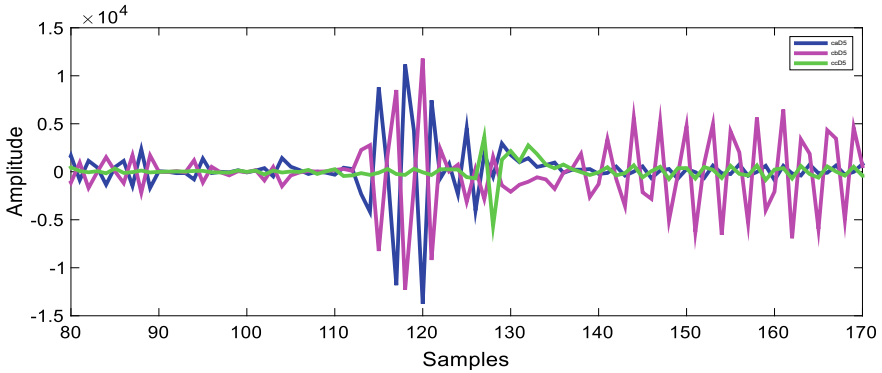


Fig. 6 RBW output when ABG fault at 0.05 s is evolved to BG fault at 0.15 s

Table 2 Results deliberating evolving faults

Case-1 (s)	F_L (km)	R_F (Ω)	R_G (Ω)	Case-2 (s)	RBW output		
					A	B	C
ABG (0.05)	125	1.25	2.25	BG (0.15)	1.1182×10^4	1.1794×10^4	3.6309×10^3
AG (0.07)	125	1.25	2.25	BG (0.17)	5.7796×10^3	7.6996×10^3	2.5154×10^3
BG (0.1)	125	1.25	2.25	ACG (0.15)	1.6437×10^4	4.0364×10^3	1.8684×10^4
ABCG (0.15)	125	1.25	2.25	CG (0.2)	5.4685×10^3	6.2966×10^3	4.5859×10^3
BCG (0.06)	125	1.25	2.25	BG (0.18)	2.8793×10^3	1.6381×10^4	1.8959×10^4

4.3 Close-End Relay Faults

The current waveform during ABG fault observed at 0.075 s is reported in Fig. 7. The waveform of RBW output is demonstrated in Fig. 8. The calculated amplitudes of RBW during several fault cases considering the modification of fault position from 5 to 9 km are presented in Table 3. The performance is found reasonably supportable as per the effects presented in Table 3 [20].

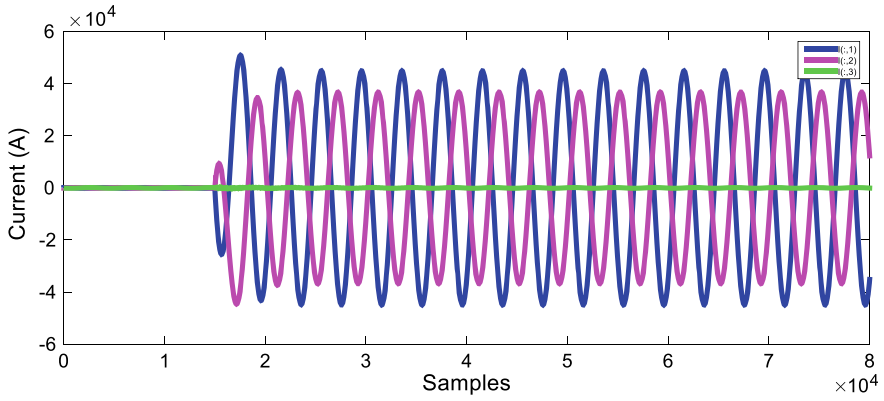


Fig. 7 ABG fault at 5 km at 0.075 s

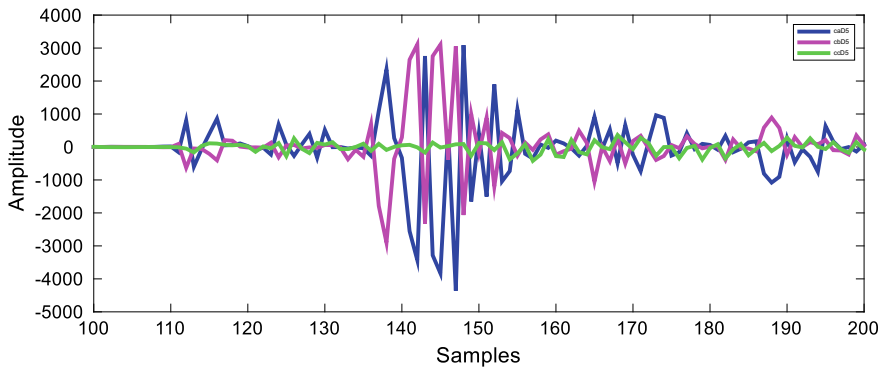


Fig. 8 RBW output for ABG fault at 5 km at 0.075 s

Table 3 Results deliberating close-end faults

Case	F_L (km)	Time (s)	RBW output		
			A	B	C
ABG	5	0.075	3.0894×10^3	3.1084×10^3	361.9971
ACG	6	0.075	4.3368×10^3	525.0509	6.4252×10^3
ABCG	7	0.075	5.3249×10^3	6.1153×10^3	7.4453×10^3
BCG	8	0.075	1.1624×10^3	2.2070×10^3	3.7330×10^3
AG	9	0.075	2.9762×10^3	1.3675×10^3	1.3764×10^3

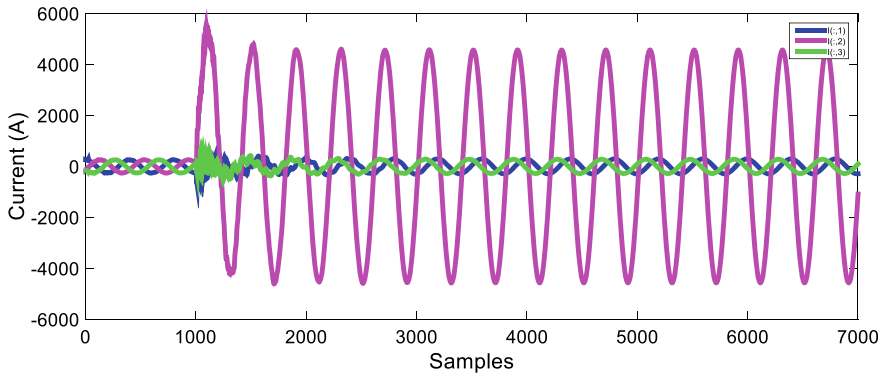


Fig. 9 BG fault at 195 km at 0.05 s

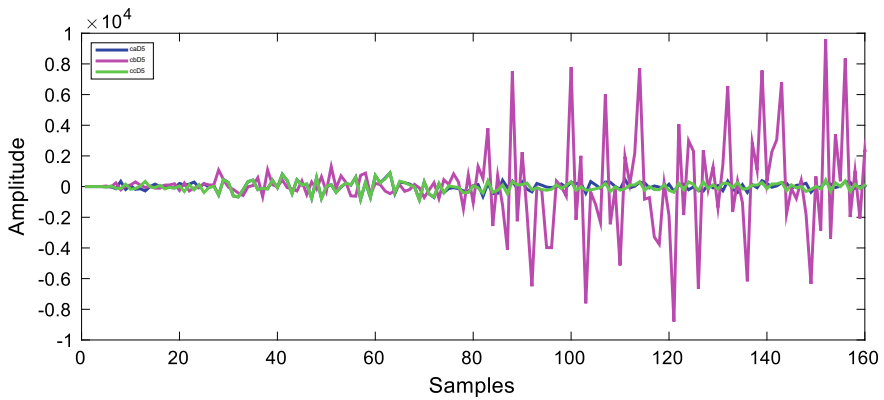


Fig. 10 RBW output for BG fault at 195 km at 0.05 s

4.4 Far-End Relay Faults

The current waveform during BG fault switched at 0.05 s is represented in Fig. 9. The waveform of obtained RBW output is demonstrated in Fig. 10. The evaluated amplitudes of RBW during numerous fault cases taking into consideration the modification of fault position from 195 to 199 km are presented in Table 4. The performance is found practically supportable as per the effects presented in Table 4 [21].

4.5 Faults at Two Different Places

The current waveform during CG fault at 147 km and AG fault at 53 km analyzed at 0.078 s is demonstrated in Fig. 11 and the obtained RBW output is presented

Table 4 Results deliberating remote-end faults

Case	F_L (km)	Time (s)	RBW output		
			A	B	C
BG	195	0.05	898.1460	9.8209×10^3	889.2847
AG	196	0.05	8.9789×10^3	422.6996	461.0553
ABG	197	0.05	2.4769×10^4	2.2789×10^4	743.6391
BCG	198	0.05	718.2602	2.0282×10^4	2.3350×10^4
ACG	199	0.05	2.2217×10^4	839.4304	2.2941×10^4

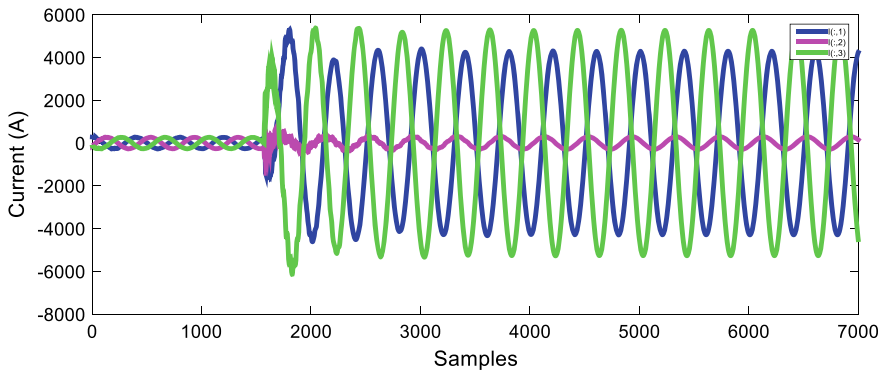


Fig. 11 CG fault at 147 km and AG fault at 53 km at 0.078 s

in Fig. 12. The calculated amplitudes of RBW during some faults are reported in Table 5. The performance of the scheme is found quite satisfactory according to the effects deliberated in Table 5 [22].

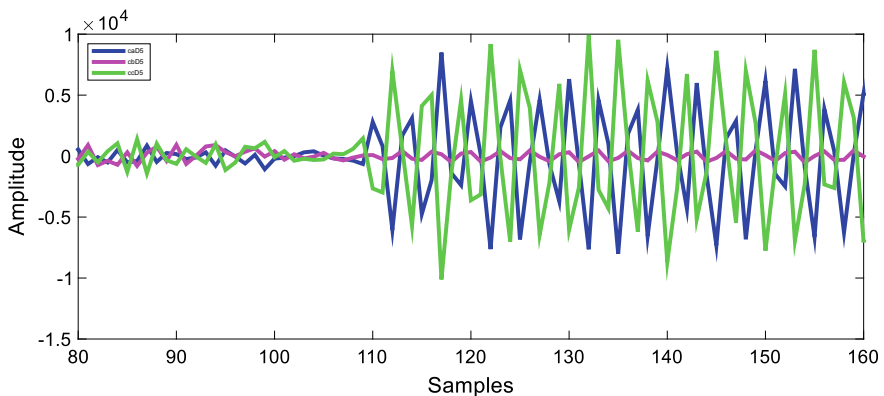


Fig. 12 RBW output for CG fault at 147 km and AG fault at 53 km at 0.078 s

Table 5 Results deliberating two different position faults

Case-1	Case-2	Time (s)	$R_F (\Omega)$	$R_G (\Omega)$	RBW Output		
					A	B	C
CG (147 km)	AG (53 km)	0.078	4.15	1.15	8.4905×10^3	1.1043×10^3	9.9061×10^3
BG (75 km)	CG (125 km)	0.078	4.15	1.15	1.1484×10^3	1.6151×10^4	7.5277×10^3
AG (110 km)	BG (90 km)	0.078	4.15	1.15	1.4534×10^4	8.1232×10^3	805.2229
CG (65 km)	BG (135 km)	0.078	4.15	1.15	1.2068×10^3	6.9895×10^3	1.6766×10^4
AG (120 km)	CG (80 km)	0.078	4.15	1.15	1.1383×10^4	1.4855×10^3	7.9151×10^3

4.6 Faults Around Series Capacitor Bank

The placement of fault initiator block is varied around the capacitor bank. The current waveform during AG fault at 50% length before capacitor bank at 0.1065 s is pointed-up in Fig. 13 and the acquired RBW plot is presented in Fig. 14. The calculated amplitudes of RBW during several faults when the placement of fault initiator block is changed with respect to capacitor bank are reported in Table 6. The operation is found extremely acceptable according to the consequences reported in Table 6.

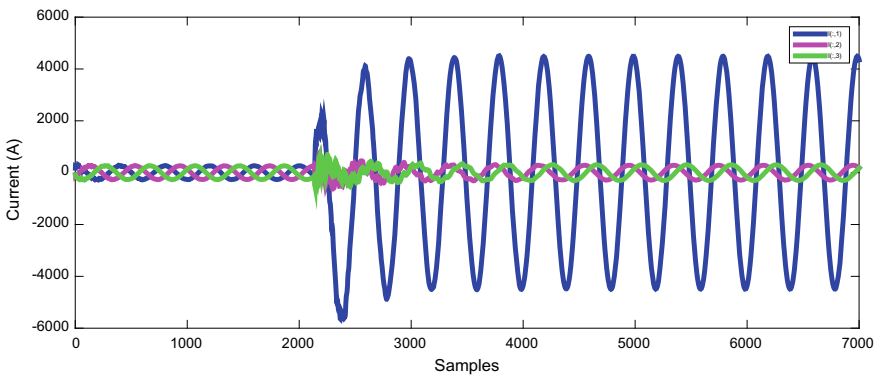


Fig. 13 AG fault at 50% length before capacitor bank at 0.1065 s

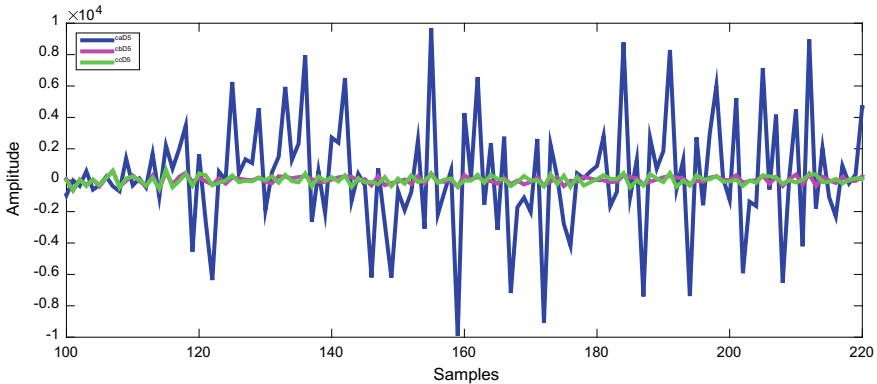


Fig. 14 RBW output for AG fault at 50% length before capacitor bank at 0.1065 s

Table 6 Results deliberating faults around capacitor bank

Case	$R_F (\Omega)$	$R_G (\Omega)$	Fault place	RBW output		
				A	B	C
AG	1.5	0.5	Before	9.6951×10^3	1.0275×10^3	1.0104×10^3
ABG	1.5	0.5	After	1.1459×10^4	1.2252×10^4	925.8613
ACG	1.5	0.5	Before	1.8472×10^4	814.2999	2.0496×10^4
BG	1.5	0.5	After	633.7007	5.9908×10^3	453.4812
ABCG	1.5	0.5	Before	5.3485×10^4	5.4692×10^4	5.3783×10^4

4.7 Different Wind Generators

The current waveform during ABCG fault at 140 km at 0.115 s amid two wind generators is reported in Fig. 15. The RBW output plot is demonstrated in Fig. 16. The deliberated amplitudes of RBW during several faults with different wind generators are presented in Table 7. The performance is found practically supportable as per the results verified in Table 7.

5 Conclusion

This paper has demonstrated an approach based on the features extracted from the fault current signals using RBW. The system and scheme are designed in MATLAB. The simulation effects specify that the scheme is steadfast as well accurate.

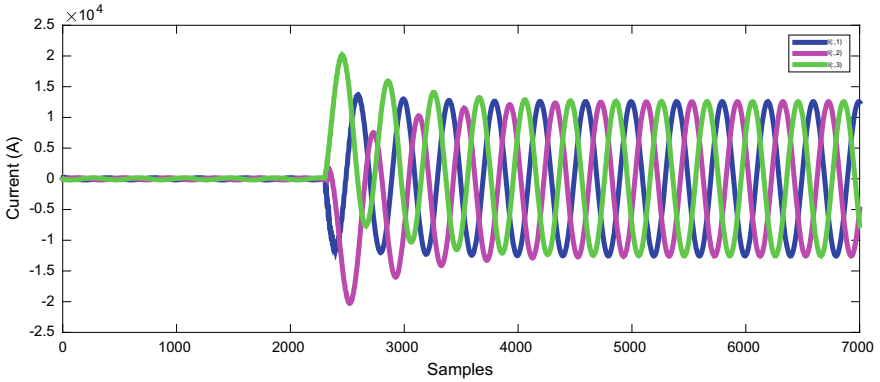


Fig. 15 ABCG fault at 140 km at 0.115 s amid two wind generators

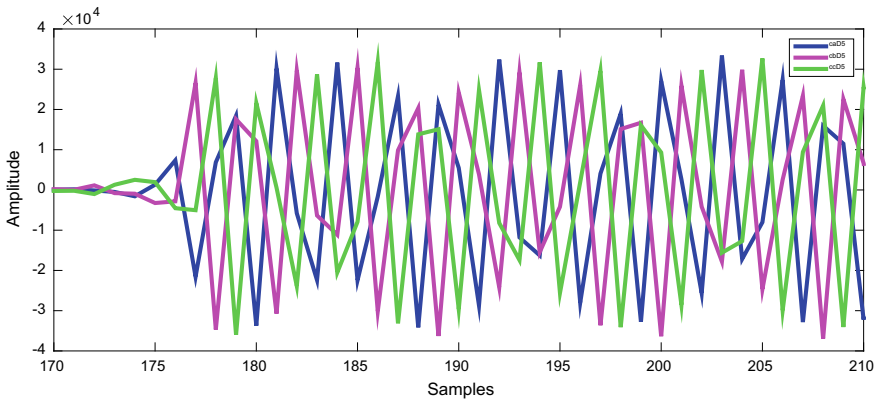


Fig. 16 RBW output for ABCG fault at 140 km at 0.115 s amid two wind generators

Table 7 Results deliberating modification in the wind generators

Case	WG	F_L (km)	RBW output		
			A	B	C
ABCG	2	140	3.3393×10^4	3.2655×10^4	3.2693×10^4
AG	4	140	1.0189×10^4	1.3752×10^3	1.3611×10^3
BCG	1	140	932.4089	2.3259×10^4	2.7525×10^4
ABG	3	140	2.5210×10^4	1.7922×10^4	412.5608
ACG	5	140	2.4566×10^4	430.5823	2.1665×10^4

References

1. E.U. Haq, H. Jianjun, K. Li et al., Improved performance of detection and classification of 3-phase transmission line faults based on discrete wavelet transform and double-channel extreme learning machine. *Electr. Eng.* (2020)
2. N. Narasimhulu, D.V.A. Kumar, M.V. Kumar, LWT based ANN with ant lion optimizer for detection and classification of high impedance faults in distribution system. *J. Electr. Eng. Technol.* **15**, 1631–1650 (2020)
3. P. Rai, N.D. Londhe, R. Raj, Fault classification in power system distribution network integrated with distributed generators using CNN. *Electr. Power Syst. Res.* 106914 (2020)
4. H.A. Jimenez, D. Guillen, R. Tapia-Olvera, G. Escobar, F. Beltran-Carbajal, An improved algorithm for fault detection and location in multi-terminal transmission lines based on wavelet correlation modes. *Electr. Power Syst. Res.* 106953 (2020)
5. B.Y. Vyas, R.P. Maheshwari, B. Das, Versatile relaying algorithm for detection and classification of fault on transmission line. *Electr. Power Syst. Res.* 106913 (2020)
6. A. Abdullah, Ultrafast transmission line fault detection using a DWT-based ANN. *IEEE Trans. Ind. Appl.* **54**(2), 1182–1193 (2018)
7. S. Affjulla, P. Tripathy, A robust fault detection and discrimination technique for transmission lines. *IEEE Trans. Smart Grid* **9**(6), 6348–6358 (2018)
8. M.K. Jena, S.R. Samantaray, B.K. Panigrahi, A new decentralized approach to wide-area back-up protection of transmission lines. *IEEE Syst. J.* **12**(4), 3161–3168 (2018)
9. E.B.M. Tayeb, O.A.A.A. Rhim, Transmission line faults detection, classification and location using artificial neural network, in *International Conference & Utility Exhibition on Power and Energy Systems: Issues and Prospects for Asia (ICUE)*, pp. 1–5 (2011)
10. A. Rahmati, R. Adhami, A fault detection and classification technique based on sequential components, in *IEEE Industry Applications Society Annual Meeting*, pp. 1–5 (2013)
11. G. Kapoor, R.N. Shaw, SWT-based fault recognition technique for six-phase transmission system, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 328–333 (2020). <https://doi.org/10.1109/gucon48875.2020.9231216>
12. G. Kapoor, R.N. Shaw, Identification of faults in parallel transmission lines using discrete Meyer wavelet, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, pp. 736–741 (2020). <https://doi.org/10.1109/iccca49541.2020.9250831>
13. G. Kapoor, R.N. Shaw, A MMG indices-based scheme for fault detection in series-compensated power transmission system, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, pp. 742–747 (2020). <https://doi.org/10.1109/iccca49541.2020.9250714>
14. R.S. Somalwar, S.G. Kadwane, R.N. Shaw, Frequency estimation by recursive least square in active islanding method for microgrid, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 116–123 (2020). <https://doi.org/10.1109/gucon48875.2020.9231206>
15. G. Kapoor, A. Mishra, R.N. Shaw, S. Ajmera, Complex faults categorization technique for series-compensated transmission lines connected with wind-turbines based on DFVHT, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 703–708 (2020). <https://doi.org/10.1109/gucon48875.2020.9231152>
16. G. Kapoor, R.N. Shaw, Protection of dual-circuit transmission line integrated with wind-park using fast Walsh-Hadamard Transform, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 480–485 (2020). <https://doi.org/10.1109/gucon48875.2020.9231092>
17. Y. Belkhier, A. Achour, R.N. Shaw, Fuzzy passivity-based voltage controller strategy of grid-connected PMSG-based wind renewable energy system, in *2020 IEEE 5th International*

- Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, pp. 210–214 (2020). <https://doi.org/10.1109/ficcca49541.2020.9250838>
18. D. Saravanan, A. Hasan, A. Singh, H. Mansoor, R.N. Shaw, Fault prediction of transformer using machine learning and DGA, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 1–5 (2020). <https://doi.org/10.1109/gucon48875.2020.9231086>
 19. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON)*, SONEPAT, India, pp. 1–4 (2020). <https://doi.org/10.1109/piicon49524.2020.9112952>
 20. D. Saravanan, D. Singh, S. Sinha, R.N. Shaw, K. Prabu, A comprehensive study of well-known maximum power point tracking techniques, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 829–837 (2020). <https://doi.org/10.1109/gucon48875.2020.9231219>
 21. M.D. Siddique R.N. Shaw et al., Switched-capacitor based seven-level triple voltage gain boost inverter (7L-TVG-BI), in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 848–852 (2020). <https://doi.org/10.1109/gucon48875.2020.9231182>
 22. Y. Belkhier, A. Achour, N. Ullah, R.N. Shaw, Modified passivity-based current controller design of permanent magnet synchronous generator for wind conversion system. *Int. J. Model. Simul.* (2020). <https://doi.org/10.1080/02286203.2020.1858226>

HWT-DCDI-Based Approach for Fault Identification in Six-Phase Power Transmission Network



Gaurav Kapoor, Pratima Walde, Rabindra Nath Shaw, and Ankush Ghosh

Abstract Determining the identification of faulty phase in six-phase power transmission network (SPPTN) is intricate in comparison with other configurations of transmission network based upon the six-phases. For this issue, the feasible faulty phase is identified with the presented scheme in this study. Haar wavelet transform (HWT) is used to the current data only as a feature extraction tool, and thereafter, a threshold method is used for deciding the faulty phase. To endorse the justification, the scheme is applied to SPPTN model implemented in MATLAB. Faults on SPPTN have been successfully identified by using the fault current measured at the sending end bus of the SPPTN.

Keywords Six-phase transmission line · Fault identification · Wavelet transform

1 Introduction

The untimely, speedy, and consistent identification of faults, due to its high probability of occurrence in TLs, is one of the mainly demanding troubles in transmission system protection. The dissemination of fault into the grid will not survive to any further extent when a fault is well-timed and appropriately detected and secluded. Various explorations have formerly been carried out the classification of faults, in the

G. Kapoor (✉)

Department of Electrical Engineering, Modi Institute of Technology, Kota, India
e-mail: gaurav.kapoor019@gmail.com

P. Walde · R. N. Shaw

Department of Electrical Engineering, Indian Institute of Technology (BHU) Varanasi, Varanasi, India
e-mail: r.n.s@ieee.org

G. Kapoor · R. N. Shaw

School of EAS, Galgotias University, Greater Noida, India

A. Ghosh

The Neotia University, Sarisha, West Bengal, India



Fig. 1 SPPTN model

literature, taking into consideration many feature extraction methods in permutation with various classification techniques. Several of those tools are: DWT and double-channel ELM [1], random forest [2], transient monitor index [3], Taylor–Fourier filter with O-splines [4]. Moreover, the fault classification and fault zone detection methodologies based on the analysis of active power and reactive power measured using PMUs are also accomplished. Additionally, other techniques are: WT, ANN and FT, WT and 2-tier MLP, covariance indices, DFT, and ED.

In this paper, a scheme is proposed for fault identification in SPPTN based on the detailed coefficients detection index (DCDI) of Haar wavelet transform (HWT).

The application of the scheme on SPPTN is presented in Sect. 2. Presented fault identification scheme is provided in Sect. 3. The effects of simulations are illustrated in Sect. 4, while Sect. 5 concludes the paper.

2 Application of the Scheme on SPPTN

The presented scheme is applied to the SPPTN of 138 kV which illustrated in Fig. 1. The model has two 60 Hz line sections of 34 km each. Numerous faults are incepted at different locations through MATLAB simulations in SIMULINK.

3 Presented Fault Identification (FI) Scheme

The block diagram of FI scheme is presented in Fig. 2.

4 Simulation Effects

The effects of simulations in view of numerous fault scenarios to bear out the implementation of the scheme are provided in this section.

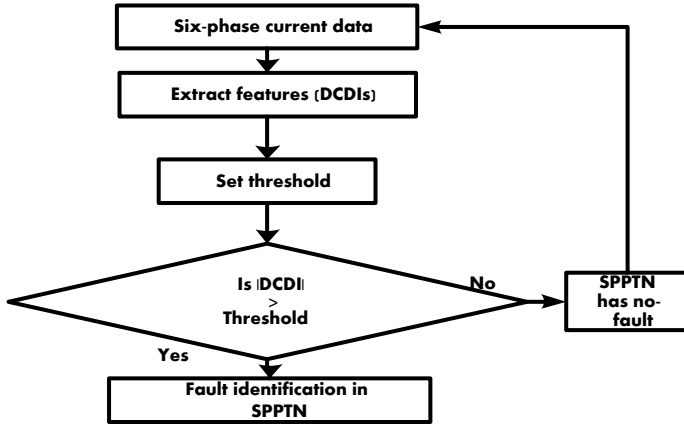


Fig. 2 Block diagram of FI scheme

4.1 Faultless Event

The current and voltage plot considering faultless event is presented in Fig. 3. The DCDIs of $\Phi A-\Phi B-\Phi C$ currents during faultless event are demonstrated in Fig. 4. The DCDIs of $\Phi D-\Phi E-\Phi F$ currents during faultless event are illustrated in Fig. 5. The HWT-DCDIs during faultless event are provided in Table 1 wherein the DCDIs are below the threshold (700) considered in this study [5, 6].

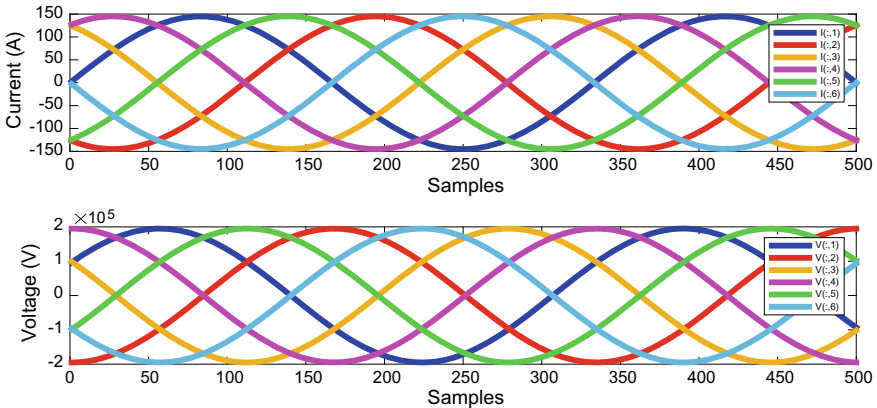


Fig. 3 Six-phase current and voltage

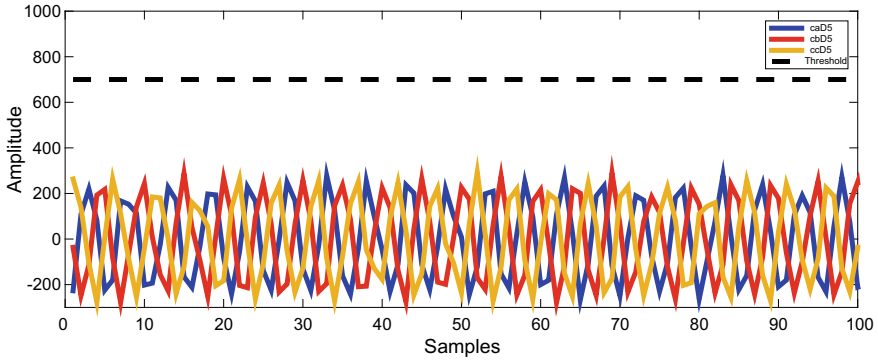


Fig. 4 DCDIs of $\Phi A-\Phi B-\Phi C$ currents during faultless event

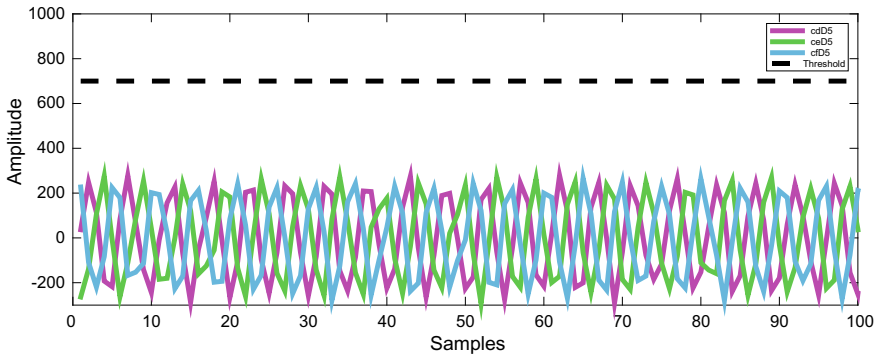


Fig. 5 DCDIs of $\Phi D-\Phi E-\Phi F$ currents during faultless event

Table 1 HWT-DCDIs during faultless event

HWT-DCDIs					
$\Phi-A$	$\Phi-B$	$\Phi-C$	$\Phi-D$	$\Phi-E$	$\Phi-F$
280.3554	290.5566	293.0473	278.9004	287.6546	266.7550

4.2 Different Faults

The current waveform of SPPTN during ABIEFG fault created at 34 km and activated at 0.05 s, $R_F = 3 \Omega$ and $R_G = 5 \Omega$ is demonstrated in Fig. 6. The DCDIs of $\Phi A-\Phi B-\Phi C$ are presented in Fig. 7. The DCDIs of $\Phi D-\Phi E-\Phi F$ are illustrated in Fig. 8. The inspection on the output of DCDI plots confirms the correct identification of ABIEFG fault. Upshots of some more faults are provided in Table 2. The activity of fault identification has been correctly performed by the scheme as examined from Table 2 [7, 8].

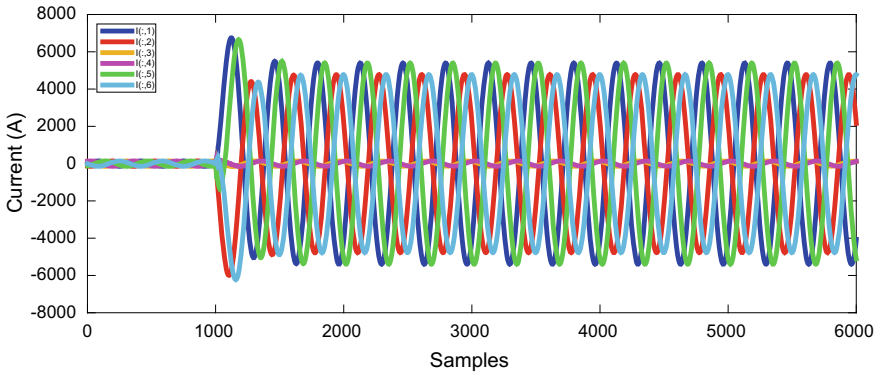


Fig. 6 ABEFG fault at 0.05 s

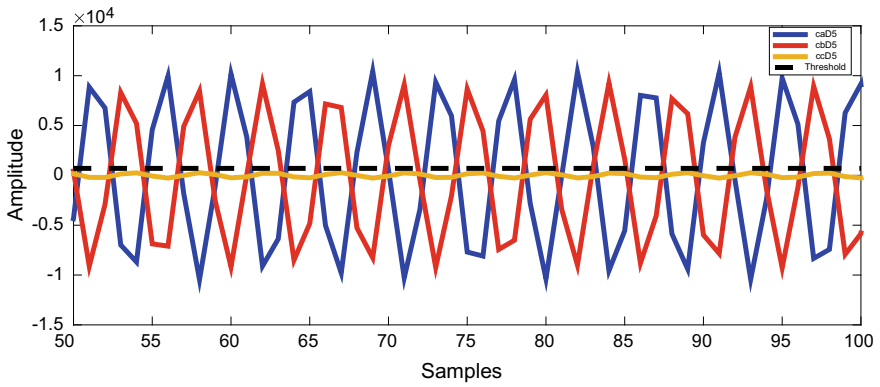


Fig. 7 DCDIs of $\Phi A-\Phi B-\Phi C$ for ABEFG fault at 0.05 s

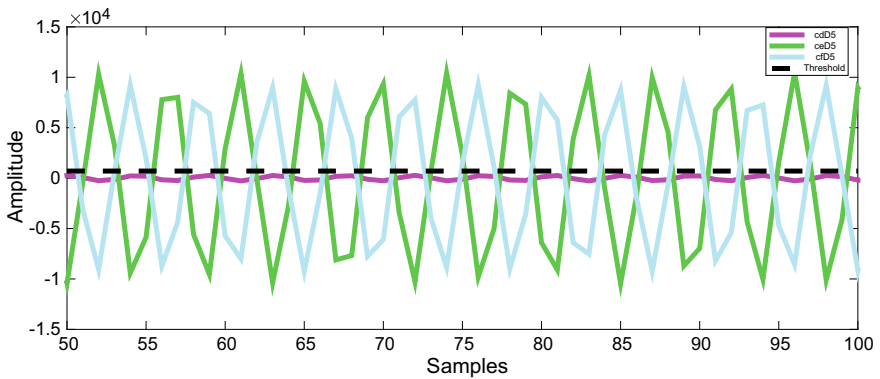


Fig. 8 DCDIs of $\Phi D-\Phi E-\Phi F$ for ABEFG fault at 0.05 s

Table 2 HWT-DCDIs during different faults

Case	HWT-DCDIs					
	Φ -A	Φ -B	Φ -C	Φ -D	Φ -E	Φ -F
ABEFG	1.0470×10^4	9.2586×10^3	276.6454	274.2060	1.0497×10^4	9.2641×10^3
ABCG	1.1739×10^4	1.1162×10^4	1.1220×10^4	273.0650	277.4110	290.3799
DEFG	273.5627	295.3993	278.2646	1.1291×10^4	1.1365×10^4	1.1095×10^4
ABG	1.0269×10^4	8.9067×10^3	275.6649	290.3128	280.9782	284.6464
FG	259.7004	261.2972	298.6755	253.1026	294.5469	6.1612×10^3

4.3 Different Fault Resistances

The current waveform of SPPTN during ABCDEG fault created at 50% distance and activated at 0.1 s, $R_F = 5 \Omega$ and $R_G = 4 \Omega$ is presented in Fig. 9. For Φ A– Φ B– Φ C, the DCDIs are presented in Fig. 10. For Φ D– Φ E– Φ F, the DCDIs are illustrated in Fig. 11. The DCDIs consequences authenticate the case of ABCDEG fault. The DCDIs calculated for various faults are presented in Table 3. The DCDIs in Table 3 evidently illustrate that despite of R_F modification; the scheme executes the FI job acceptably [9, 10].

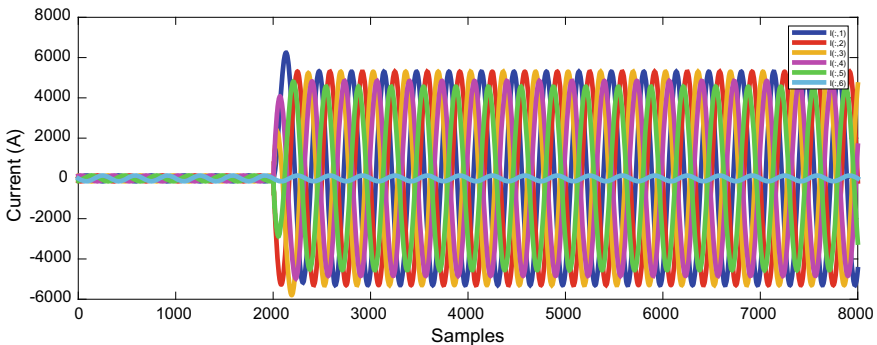


Fig. 9 ABCDEG fault at 0.1 s

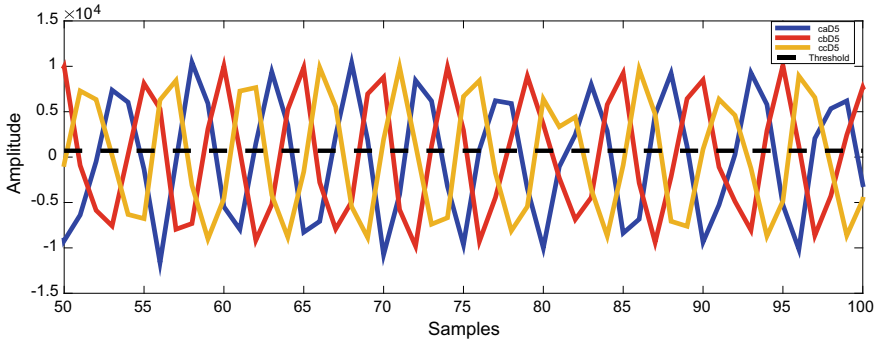


Fig. 10 DCDIs of $\Phi A-\Phi B-\Phi C$ for ABCDEG fault at 0.1 s

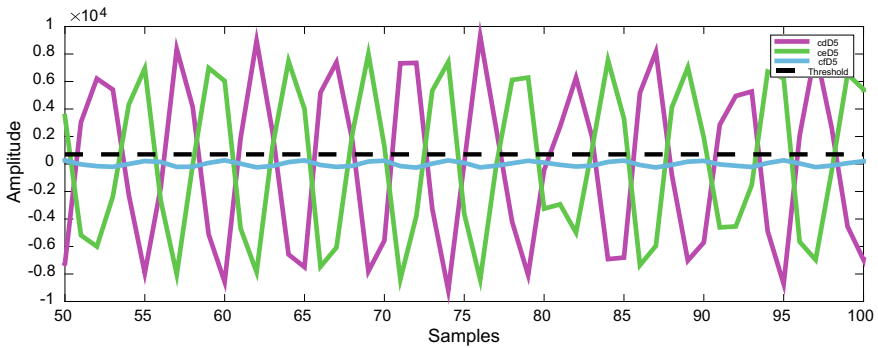


Fig. 11 DCDIs of $\Phi D-\Phi E-\Phi F$ for ABCDEG fault at 0.1 s

Table 3 HWT-DCDIs during different fault resistances

Case	HWT-DCDIs					
	$\Phi-A$	$\Phi-B$	$\Phi-C$	$\Phi-D$	$\Phi-E$	$\Phi-F$
ABCDEG (5 Ω)	1.1294×10^4	1.0373×10^4	1.0921×10^4	9.2919×10^3	1.0018×10^4	275.9218
CDEFG (25 Ω)	279.7904	282.8992	3.1928×10^3	4.0911×10^3	4.6097×10^3	4.2293×10^3
ABFG (45 Ω)	2.3720×10^3	2.5432×10^3	310.1047	278.4545	276.6609	2.4664×10^3
CDG (55 Ω)	311.9066	289.9350	2.0346×10^3	2.0277×10^3	284.4262	298.4495
AG (75 Ω)	1.5769×10^3	284.0488	285.3738	286.7058	282.5331	282.6707

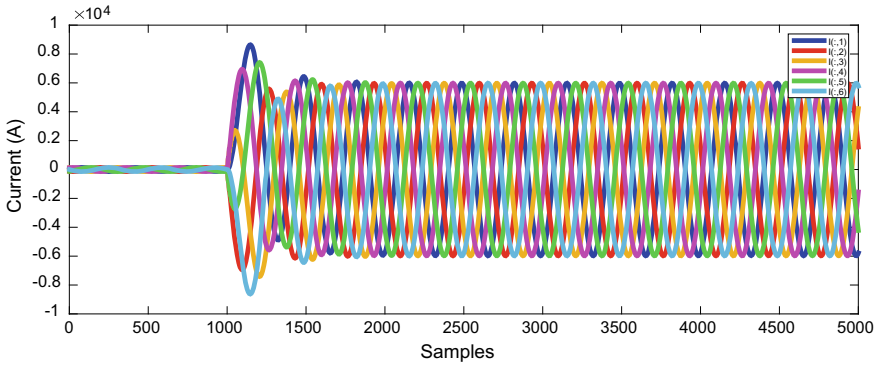


Fig. 12 ABCDEFG fault at 0.05 s

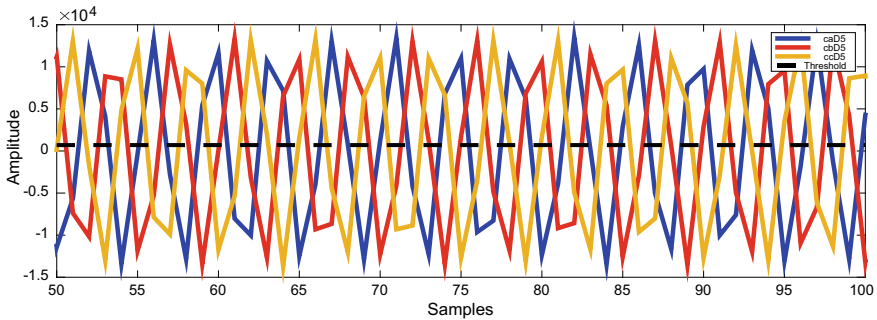


Fig. 13 DCDIs of $\Phi A-\Phi B-\Phi C$ for ABCDEFG fault at 0.05 s

4.4 Different Fault Switching Time

The current waveform of SPPTN during ABCDEFG fault created at 50% length and initiated at 0.05 s, $R_F = 1.85 \Omega$ and $R_G = 3.25 \Omega$ is displayed in Fig. 12. The DCDIs of $\Phi A-\Phi B-\Phi C$ are presented in Fig. 13. The DCDIs of $\Phi D-\Phi E-\Phi F$ are illustrated in Fig. 14. The scrutiny on the output of DCDI plots approves the accurate classification of ABCDEFG fault. Consequences of some more faults are provided in Table 4. The action of FI has been suitably performed by the scheme as observed from Table 4 [11, 12].

4.5 Different Fault Locations

The current waveform of SPPTN during ABCEFG fault created at 20 km and incepted at 0.07 s, $R_F = 3.15 \Omega$ and $R_G = 2.25 \Omega$ is illustrated in Fig. 15. For $\Phi A-\Phi B-\Phi C$,

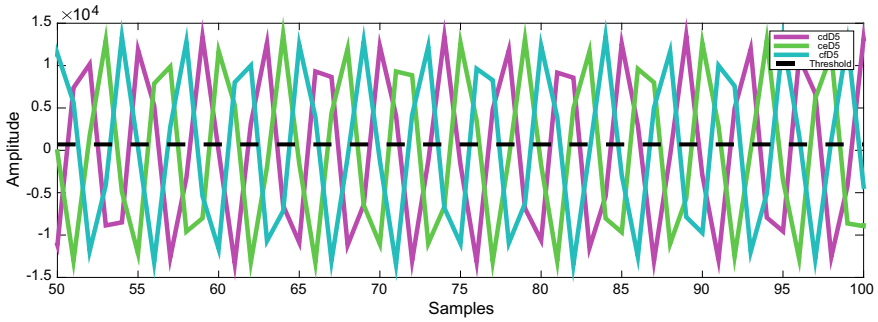


Fig. 14 DCDIs of Φ D- Φ E- Φ F for ABCDEFG fault at 0.05 s

Table 4 HWT-DCDIs during different FST

Case (s)	HWT-DCDIs					
	Φ -A	Φ -B	Φ -C	Φ -D	Φ -E	Φ -F
ABCDEFG (0.05)	1.3664×10^4	1.3445×10^4	1.3244×10^4	1.3788×10^4	1.4043×10^4	1.3460×10^4
ABDEFG (0.1)	1.1319×10^4	9.6204×10^3	289.8606	1.2602×10^4	1.3891×10^4	1.2495×10^4
BCDEG (0.06)	259.3889	1.0269×10^4	9.4322×10^3	1.0282×10^4	9.5575×10^3	259.0332
ADFG (0.14)	5.9531×10^3	263.6839	261.7073	9.1980×10^3	260.3931	1.0343×10^4
BCG (0.08)	269.9758	1.0664×10^4	9.0870×10^3	269.3678	297.7781	274.7584

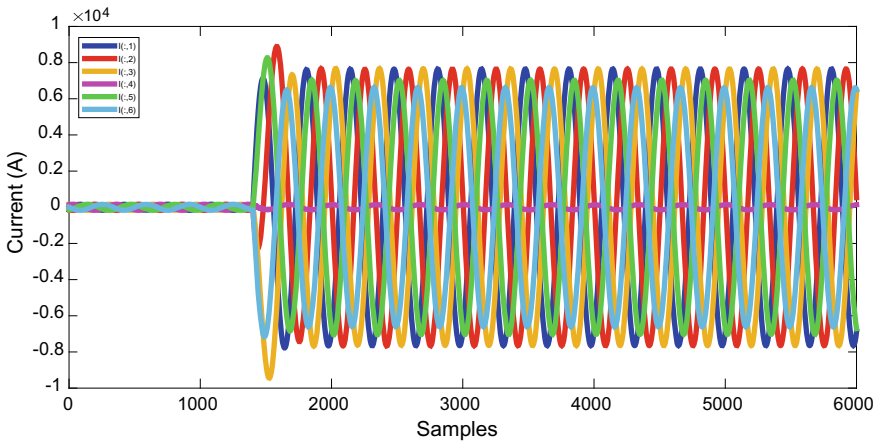


Fig. 15 ABCEFG fault at 20 km

the DCDIs are presented in Fig. 16. For $\Phi D-\Phi E-\Phi F$, the DCDIs are demonstrated in Fig. 17. The DCDIs effects validate the case of ABCEFG fault. The DCDIs evaluated for some more faults are presented in Table 5. The DCDIs in Table 5 obviously show that regardless of F_L modification; the scheme carries out the fault identification job adequately [13–16].

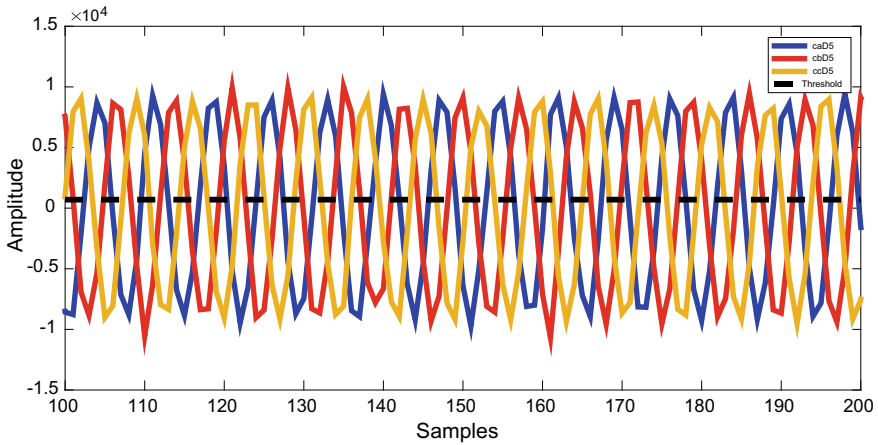


Fig. 16 DCDIs of $\Phi A-\Phi B-\Phi C$ for ABCEFG fault at 20 km

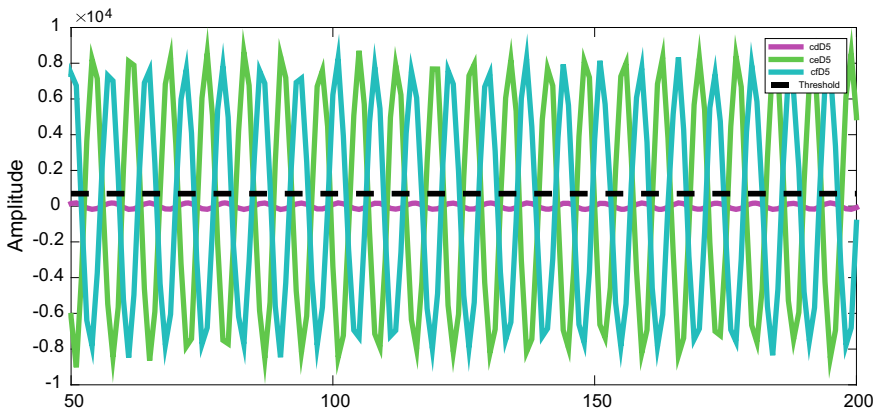


Fig. 17 DCDIs of $\Phi D-\Phi E-\Phi F$ for ABCEFG fault at 20 km

Table 5 HWT-DCDIs during different fault locations

Case (km)	HWT-DCDIs					
	Φ -A	Φ -B	Φ -C	Φ -D	Φ -E	Φ -F
ABCEFG (20)	9.4606×10^3	1.0014×10^4	9.9237×10^3	257.9048	8.8767×10^3	8.3781×10^3
ACDEG (25)	9.0266×10^3	282.4831	1.0141×10^4	1.0060×10^4	8.8503×10^3	272.4168
BCDG (30)	273.8098	1.1044×10^4	9.6034×10^3	6.7244×10^3	292.4995	264.6059
ABG (35)	9.6865×10^3	8.7104×10^3	255.7563	306.2751	276.3598	278.8764
DG (40)	311.3046	270.4703	279.1540	6.5079×10^3	268.0780	351.8673

5 Conclusion

This paper has demonstrated a substitute scheme based on threshold for fault identification for SPPTN. The presented scheme utilizes the fault data (six-phase currents) assessed only from a one evaluating spot (sending end bus) of the SPPTN. Characteristics withdrawal method was utilized, and the fault identification task was operated with threshold-based scheme.

For the substantiation of the scheme, SPPTN was employed. It can be said as per the results that the presented scheme can detect the faulty phase faultlessly.

References

1. E.U. Haq, H. Jianjun, K. Li, et al., Improved performance of detection and classification of 3-phase transmission line faults based on discrete wavelet transform and double-channel extreme learning machine. *Electr. Eng.* (2020)
2. H. Okumus, F.M. Nuroglu, A random forest-based approach for fault location detection in distribution systems. *Electr. Eng.* (2020)
3. S. Biswas, P.K. Nayak, A fault detection and classification scheme for unified power flow controller compensated transmission lines connecting wind farms. *IEEE Syst. J.*, 1–10 (2020)
4. D. Guillen, J.A. de la O Serna, A. Zamora-Mendez, M. Paternina, F. Salinas, Taylor-Fourier filter bank Implemented with O-splines for the detection and classification of faults. *IEEE Trans. Ind. Inform.* (2020)
5. Y. Belkhier, A. Achour, R.N. Shaw, Fuzzy passivity-based voltage controller strategy of grid-connected PMSG-based wind renewable energy system, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, pp. 210–214 (2020). <https://doi.org/10.1109/iccca49541.2020.9250838>
6. G. Kapoor, R.N. Shaw, Identification of faults in parallel transmission lines using discrete Meyer wavelet, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, pp. 736–741 (2020). <https://doi.org/10.1109/iccca49541.2020.9250831>

7. G. Kapoor, R.N. Shaw, A MMG indices-based scheme for fault detection in series-compensated power transmission system, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, pp. 742–747 (2020). <https://doi.org/10.1109/iccca49541.2020.9250714>
8. D. Saravanan, A. Hasan, A. Singh, H. Mansoor, R.N. Shaw, Fault prediction of transformer using machine learning and DGA, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 1–5 (2020). <https://doi.org/10.1109/gucon48875.2020.9231086>
9. G. Kapoor, R.N. Shaw, Protection of dual-circuit transmission line integrated with wind-park using fast Walsh-Hadamard transform, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 480–485 (2020). <https://doi.org/10.1109/gucon48875.2020.9231092>
10. G. Kapoor, A. Mishra, R.N. Shaw, S. Ajmera, Complex faults categorization technique for series-compensated transmission lines connected with wind-turbines based on DFWHT, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 703–708 (2020). <https://doi.org/10.1109/gucon48875.2020.9231152>
11. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON)*, SONEPAT, India, pp. 1–4 (2020). <https://doi.org/10.1109/piicon49524.2020.9112952>
12. G. Kapoor, R.N. Shaw, SWT-based fault recognition technique for six-phase transmission system, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 328–333 (2020). <https://doi.org/10.1109/gucon48875.2020.9231216>
13. D. Saravanan, D. Singh, S. Sinha, R.N. Shaw, K. Prabu, A comprehensive study of well-known maximum power point tracking techniques, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 829–837 (2020). <https://doi.org/10.1109/gucon48875.2020.9231219>
14. R.S. Somalwar, S.G. Kadwane, R.N. Shaw, Frequency estimation by recursive least square in active islanding method for microgrid, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 116–123 (2020). <https://doi.org/10.1109/gucon48875.2020.9231206>
15. M.D. Siddique, R.N. Shaw et al., Switched-capacitor based seven-level triple voltage gain boost inverter (7L-TVG-BI), in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 848–852 (2020). <https://doi.org/10.1109/gucon48875.2020.9231182>
16. Y. Belkhier, A. Achour, N. Ullah, R.N. Shaw, Modified passivity-based current controller design of permanent magnet synchronous generator for wind conversion system. *Int. J. Model. Simul.* (2020). <https://doi.org/10.1080/02286203.2020.1858226>
17. M. Abasi, A. Saffarian, M. Joorabian, S.S. Ghodrattollah, Fault classification and fault area detection in GUPFC-compensated double-circuit transmission lines based on the analysis of active and reactive powers measured by PMUs. *Measurement* **169** (2021)
18. H.V. Gururaja Rao, N. Prabhu, R.C. Mala, Wavelet transform-based protection of transmission line incorporating SSSC with energy storage device. *Electr. Eng.* **102**, 1593–1604 (2020)
19. O.O. Kalu, T.C. Madueme, Application of artificial neural network (ANN) to enhance power systems protection: a case of the Nigerian 330 kV transmission line. *Electr. Eng.* **100**, 1467–1479 (2018)
20. M.N. Mahmud, M.N. Ibrahim, M.K. Osman et al., A robust transmission line fault classification scheme using class-dependent feature and 2-Tier multilayer perceptron network. *Electr. Eng.* **100**, 607–623 (2018)

21. M.H.H. Musa, Z. He, L. Fu, Y. Deng, A covariance indices based method for fault detection and classification in a power transmission system during power swing. *Int. J. Electr. Power Energy Syst.* **105**, 581–591 (2019)
22. ChD Prasad, P.K. Nayak, A DFT-ED based approach for detection and classification of faults in electric power transmission networks. *Ain Shams Eng. J.* **10**(1), 171–178 (2019)

Graphene Properties and Its Utility for High-Frequency Antennas



Subodh Kumar Tripathi and Rohit Tripathi

Abstract Graphene is one of the latest materials, and it has shown a new path for different fields of engineering and technology. A property of graphene especially surface conductivity has been explored extensively. Surface conductivity mathematical model of graphene is discussed, and effect of chemical voltage is explored to analyze the tuning behavior of graphene. MATLAB codes are utilized to describe the surface conductivity of graphene, and its utilization is explained in electromagnetic applications like high-frequency antenna design and future wireless sensor networks. Study reveals that graphene may be game changer material for future high-speed wireless communication and future sensor networks.

Keywords Graphene · Surface conductivity · High frequency · Chemical voltage

1 Introduction

Graphene is a latest research material made from graphite that has attracted the research society due to its uniqueness [1–3]. The most important issue in utilizing graphene substance for several electronics and communication applications is to mathematically characterize the material that would give useful properties at higher frequencies, i.e., in terahertz range. Higher frequency range especially in terahertz range has many possibilities in applications like imaging, spectroscopy, sensing and detection [4].

Diamond and graphite signified the only recognized allotropes of carbon for an extended time when the condition altered with the innovation of fullerenes [5]. A fullerene (0D) is also an allotrope of carbon with a similar arrangement to graphite. Figure 1 portrays the carbon allotropic arrangements in dissimilar forms [6].

S. K. Tripathi (✉)

Department of Electronics and Communication Engineering, Meerut Institute of Engineering & Technology, Meerut, UP, India

R. Tripathi

Department of Electronics and Communication Engineering, Galgotias University, Greater Noida, UP, India

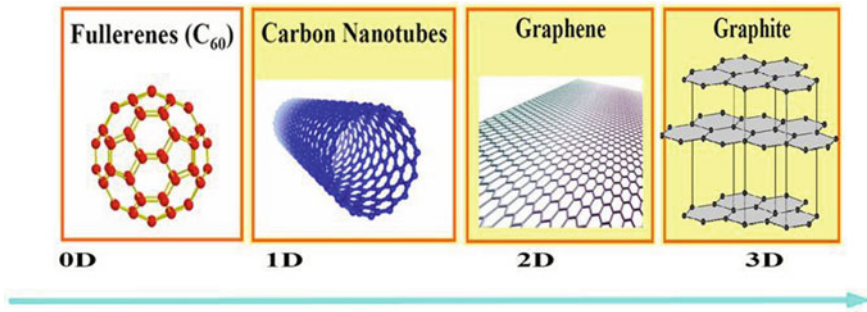


Fig. 1 Allotropic structures of the carbon [6]

Saito and Charlier [7, 8] considered that CNTs are attained by arranging graphene in proper direction. Hence, CNTs have only hexagons and may be understood as one-dimensional (1D) substances. CNTs can be of two types, i.e., single-walled nanotube and multi-walled nanotube, as shown in Fig. 2a in Fig. 2b [9].

2 Properties of Substance Graphene

A monographite layer is a plentiful substance which is an allotrope of carbon that is made from bonded carbon atoms systematized into a hexagonal lattice. SP₂ hybridization and very thin atomic thickness approximately 0.345 nm makes graphene very special. Strength, electricity, heat conduction and optical conductivity make it exceptional. 2D existence of graphene is because carbon-to-carbon bonds are so small and robust that they stop thermal variations from disrupting it. Graphene has many exceptional properties that are discussed in the next section.

Graphene is a no-overlap material with both holes and electrons as charge carriers with very extraordinary electrical conductivity. Electronic mobility of graphene is very high, and practically maximum value reported is 15,000 cm²/v.s, and maximum theoretical value is 20,000 cm²/v.s. It is considered that electronic mobility of graphene is restricted by scattering of graphene acoustic photons. It is also considered that graphene electrons act very much like photons in their mobility due to their lack of mass. Another important and exceptional property of graphene is mechanical strength. It is also reported till date that graphene is the toughest material discovered till date. Graphene strength reported till date is 130,000,000,000 Pascals or 130 Giga Pascals (G Pa) compared to 0.4 G Pa of structural steel. In respect of weight, graphene is considered light in weight approximately 0.77 mgm/m², 1 m² paper is considered as 1000 times more weighted than graphene. As per availability of graphene, thickness ranges from approximately 1 nm to 8 nm, spring constant ranges from 1 to 5 N/m and Young's modulus that is the property of elasticity being able to retain its initial state after strain.

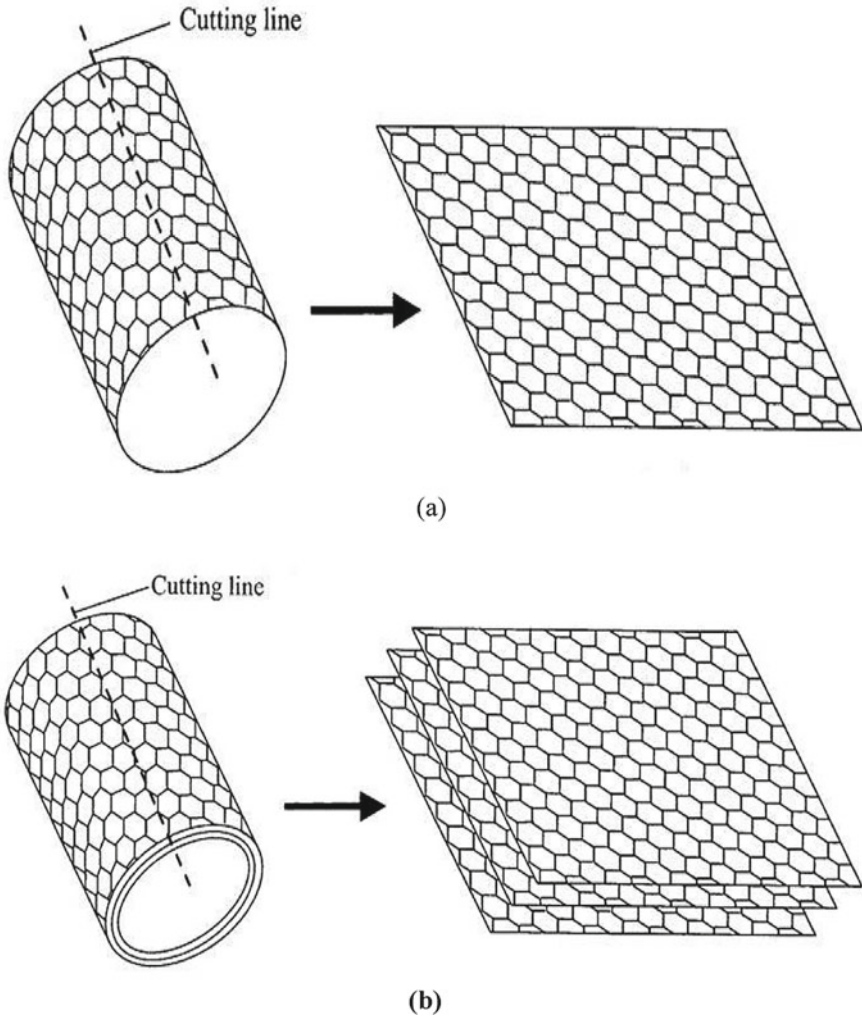


Fig. 2 **a** Single carbon nanotube made up of a single graphene sheet, **b** multi carbon nanotubes composed of graphite [9]

Graphene being very important material shows quite good properties in optical range of frequencies, i.e., 3 terahertz 3 pentahertz. Optical conductivity of the graphene is governed by following equations:

$$\sigma(\omega) = \frac{e^2}{4h} \frac{\sinh\left(\frac{h\omega}{2KT}\right)}{\cosh\left(\frac{E_f}{KT}\right) + \cosh\left(\frac{h\omega}{2KT}\right)} \tag{1}$$

$$\sigma(\omega) = \frac{e^2}{\pi h^2} \frac{2K\tau}{\tau^{-1} - j\omega} \ln\left(2 \cosh\left(\frac{E_f}{2KT}\right)\right) \quad (2)$$

where E_f is Fermi Energy, K is Boltzmann constant, T is temperature, h is Plank's constant, ω is angular frequency and τ is relaxation time. Equations 1 and 2 indicate that conductivity of graphene in optical frequency range depends upon several parameters such as Plank's constant, Fermi energy, temperature, relaxation time. It is significant to note that by changing the relaxation time and Fermi energy, conductivity can be made variable and so tuning property can be achieved in graphene. For a given values of $E_f = 89.8$ meV and $\tau^{-1} = 14.3 \times 10^{12} \text{ S}^{-1}$, conductivity value may be obtained approximately in the range of 0.5×10^6 – $2.2 \times 10^6 \text{ } \Omega^{-1} \text{ m}^{-1}$ and also For a given values of $E_f = 192.1$ meV and $\tau^{-1} = 25.4 \times 10^{12} \text{ S}^{-1}$, conductivity value may be obtained approximately in the range of 1.4×10^6 – $2.6 \times 10^6 \text{ } \Omega^{-1} \text{ m}^{-1}$.

Permittivity and optical conductivity of graphene in the optical frequency range can be related as

$$\varepsilon(\omega) = 5.5\varepsilon_0 + J \frac{\sigma(\omega)}{\omega d} \quad (3)$$

where ε_0 permittivity of free space, d is the thickness of the graphene that is approximately 0.38 nm. Amount $5.5\varepsilon_0$ is considered as background permittivity.

Graphene's surface conductivity is well determined using Kubo's formula. The expression of surface conductivity of an infinite graphene sheet is mentioned in Eq. (1), comprises two parts, intraband as a first term and interband as a second term [10]. The graphene conductivity model is specified by [11]

$$\sigma(i\omega, \mu_c, T, \Gamma) = \frac{q_e^2 (\omega + 2i\Gamma)}{\pi h^2} \left[\frac{q_e^2}{(\omega + 2i\Gamma)^2} \int_0^\infty \xi \left(\frac{\partial f_d(\xi)}{\partial \xi} - \frac{\partial f_d(-\xi)}{\partial \xi} \right) d\xi - \int_0^\infty \frac{f_d(-\xi) - f_d(\xi)}{(\omega + 2i\Gamma)^2 - \left(\frac{4\xi}{h}\right)^2} d\xi \right] \quad (4)$$

Surface conductivity of the graphene sheet is very important considering the design and modeling of devices at higher frequencies especially at optical and terahertz frequency. We can observe from given expressions (Eqs. 1–4) that conductivity of the graphene is variable and depends upon various parameters. This is important that conductivity of the graphene can be altered dynamically by changing D.C. bias voltage by means of chemical potential and fixing relaxation time. This tunable and variable conductivity can be very useful for the design various tunable devices. Graphene's surface conductivity is analyzed mathematically extensively and is coded using MATLAB (Matrix laboratory). The values of various parameters are selected according, and chemical voltage is made variable parameters whose value is varied from 0 to 1 eV. Frequency of interest is varied from 0 to 10 terahertz. Real and imaginary parts of surface conductivity of the graphene with respect to frequency are shown in Fig. 3. In conductivity, inductive nature can be observed from Fig. 3 that permits graphene to assist surface plasmons [12].

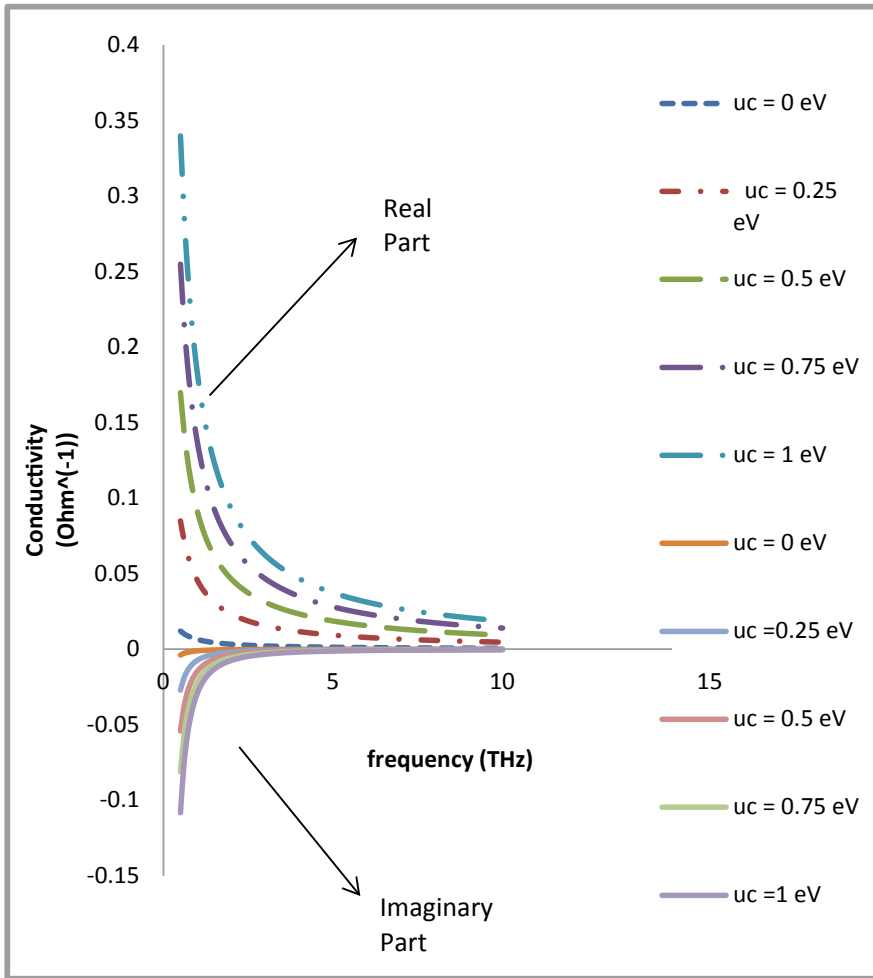


Fig. 3 Real and imaginary parts of surface conductivity of graphene

3 Utilization and Applications of Substance Graphene

Substance graphene has become one of the very important latest research materials that find its application in several fields. Graphene is being used in wireless communication especially in the design of terahertz antenna due to its conductivity in terahertz frequency and so it can be used to increase the data rate for short-distance indoor communications [13–15]. Graphene also finds applications in electronics and photonics devices and circuits to improve the performance due to lightweight and mechanical strength properties. Graphene is also reported as a very useful material for the design different types of sensors [16]. It is also reported that graphene may

be useful in several other applications like in solar cells and medical fields due to exceptional properties of graphene [17, 18].

4 Results and Discussions

Surface conductivity of the graphene is modeled mathematically and plotted versus frequency considering the effects of chemical voltage as depicted in Fig. 3. Both real and imaginary parts are shown there. We may easily observe the effects of chemical voltage upon the conductivity of the graphene sheet. This plot shows that conductivity of the graphene sheet is dependent and may be used as dynamic parameter for designing of high frequency specially terahertz antenna. This electronically tunable frequency tunable material may be very beneficial for making reconfigurable high frequency antenna and fast and better wireless communication. These frequency-dependent antennas may also be useful in wireless sensor networks of future wireless communication in nano scale.

5 Challenges and Difficulties in Using Substance Graphene

One of the major challenges of the graphene industry is to produce volume production in the coming years. The focus required is to be on substance consistency and cost of graphene production. Cost is one of very significant factor in itself. The problem lies in achieving and sustaining the superiority in producing bulk quantities of graphene.

Defects on the graphene monolayer extensively affect the electrical property, impermeability, thermal property and transparency and all other important properties that are subjected to graphene. At present, two important methods are used to produce graphene, i.e., top down and bottom up. In bottom up, monolayer is created by assembling carbon atoms, and chemical vapor deposition (CVD) process is utilized in this method, which permits to make a graphene monolayer upon a nickel or copper or metallic substrate. If we have to put graphene upon any substrate, in that case we will have to transfer it from the metal, this process may cause different defects upon monolayers like cracks, crinkles and holes. These defects may cause fall in the qualities and overall performance of the graphene. This problem is the major one concerning different electronics projects made using graphene. These projects may include electronic devices like smart mobile phone, nano-wireless sensors, etc.

Graphite exfoliation process for extracting graphene refers to top-down method.

Graphite consists of various graphene monolayers that are stuck to each other. This top-down method permits to break all stuck layers to produce single layer of graphene. Top-down method necessarily indicates about starting from a big one and ending with a small one. Sticky-type method was utilized first to extract single layer of graphene. As of now exfoliation method is best available method which is defect less to produce graphene.

6 Conclusion

In this paper, graphene, its properties, applications and future challenges are studied and reported. It is observed that graphene has been very significant and useful material since the invention of the material. Some important properties of the graphene are discussed and explained. In specialty, graphene's surface conductivity is mathematically modeled and shown the use in higher frequency, i.e., terahertz frequency. Tunable property can be very useful in designing of reconfigurable antenna in very small scale useful in future high-speed wireless indoor communication and in wireless sensor networks. In addition, other graphene's applications like in electronics, photonics and medical fields are discussed. Constraint and future challenges in utilization of graphene are also discussed here. It is observed that major challenge in graphene's applications lies on large and volume production of the substance graphene and its initial manufacturing cost.

Acknowledgements The authors would like to convey their gratitude for the support provided by Meerut Institute of Engineering & Technology, Meerut, UP, India, IIT Indore, India and Thapar University, Patiala, Punjab, India, I.K.G. Punjab Technical University, Jalandhar, Punjab, India. The authors convey their thanks to Dr. Mukesh Kumar, Associate professor, IIT, Indore, India for his kind guidance and support.

References

1. K. Geim, K.S. Novoselov, The rise of graphene. *Nat. Mater.* **6**, 183–191 (2007)
2. P. Blake, E.W. Hill1, A.H.C. Neto, K.S. Novoselov, D. Jiang, R. Yang, T.J. Booth, A.K. Geim, Making grapheme visible. *Appl. Phys. Lett.* **91**, 063124 (2007)
3. C. Stampfer, J. Güttinger, F. Molitor, D. Graf, T. Ihn, K. Ensslin, Tunable coulomb blockade in nano structured graphene. *Appl. Phys. Lett.* **92**, (2008)
4. I.F. Akyildiz, M. Jornet, C. Han, Terahertz band: next frontier for wireless communication. *Phys. Commun.* **12**, 16–32 (2014)
5. H.W. Kroto, J.R. Heath, R.E. Smalley, C60: Buckminsterfullerene. *Nature Publishing Group*, Nov. 318: 162–163 (1985)
6. G.G. Hoffmann, G. de With, and J. Loos, Micro-Raman and tip-enhanced Raman spectroscopy of carbon allotropes. *Macromol. symposia* **265**(1), 1–11 (2008)
7. R. Saito, G. Dresselhaus, M.S. Dresselhaus, *Physical Properties of Carbon Nanotubes* (Imperial College Press, London, 1998)
8. J.-C. Charlier, X. Blase, S. Roche, Electronic and transport properties of nanotubes. *Rev. Mod. Phys.* **79**(2), 677 (2007)
9. (Patent) The original graphene patent and the rift between industry and academia By Marko Spasenovic on January 15, 2013 <http://www.graphenetracker.com/the-original-graphene-patent>
10. G.W. Hanson, Dyadic Green's functions and guided surface waves for a surface conductivity model of graphene. *J. Appl. Phys.* **103**(6), 064302 (2008). <https://doi.org/10.1063/1.2891452>
11. J.M. Jornet, I.F. Akyildiz, Graphene based nano antennas for electromagnetic nano communications in the terahertz band, in *Proceedings of the Fourth European Conference on Antennas and Propagation* (Barcelona, Spain, 2010), pp. 1–5. <https://doi.org/10.1016/j.nancom.2010.04.001>

12. I. Das, R.N. Shaw, S. Das, Location-based and multipath routing performance analysis for energy consumption in wireless sensor networks, in *Innovations in Electrical and Electronic Engineering*. Lecture Notes in Electrical Engineering, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh, vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_59
13. S.K. Tripathi, A. Kumar, Performance comparison of graphene terahertz antenna and copper terahertz antenna, in *Advances in Computational Intelligence and Communication Technology*. Advances in Intelligent Systems and Computing, ed. by X.Z. Gao, S. Tiwari, M. Trivedi, K. Mishra, vol. 1086 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-1275-9_40
14. S.K. Tripathi, M. Kumar, A. Kumar, Graphene based tunable and wideband terahertz antenna for wireless network communication. *Wireless Netw.* **25**, 4371–4381 (2019). <https://doi.org/10.1007/s11276-019-02101-8>
15. Subodh Kumar Tripathi & Ajay Kumar, High gain miniaturised photonic band gap terahertz antenna for size-limited applications. *Aust. J. Electr. Electron. Eng.* **16**(2), 74–80 (2019). <https://doi.org/10.1080/1448837X.2019.1602944>
16. Y. Xu, Z. Guo, H. Chen, Y. Yuan, J. Lou, X. Lin, H. ao, H. Chen, B. Yu, In-plane and tunneling pressure sensors based on graphene/hexagonal boron nitride heterostructure. *Appl. Phys. Lett.* **99**(13), 133109 (2011). <https://doi.org/10.1063/1.3643899>
17. M. Orcutt, Research Hints at Graphene’s Photovoltaic Potential, Newly observed properties mean graphene could be a highly efficient converter of light to electric power. MIT. March 1, 2013
18. A. Fabbro et al., Graphene-Based Interfaces do not Alter Target Nerve Cells. *ACS Nano* (2016). <https://doi.org/10.1021/acsnano.5b05647>

Comparative Analysis of MOSFET, FINFET and GAAFET Devices Using Different Substrate and Gate Oxide Materials



Ritik Koul, Mukul Yadav, and Rajeshwari Pandey

Abstract This work investigates the efficacy of the different substrate and oxide materials for the three devices—MOSFET, FINFET and GAAFET. In this respect, the analog performance parameters of these devices such as on-current (I_{on}), off-current (I_{off}), current switching ratio (I_{on}/I_{off}) and sub-threshold swing (SS) are examined for the different substrate and oxide materials used in the modeling of the devices. Different substrate materials used in this work are Si, SiGe, GaAs and SiC₃C, and different oxide materials used are SiO₂ and HfO₂. The simulation is done using the COGENDA VisualTCAD tool.

Keywords Gate electrostatic control · Current switching ratio · Sub-threshold swing · MOSFET · FINFET · GAAFET

1 Introduction

There has been always a need for a compact, fast and energy-efficient device which led to the continuous scaling down of the semiconductor devices like MOSFET. But, short channel effects such as hot electron effect, carrier velocity saturation, drain induced barrier lowering, leakage current and sub-threshold swing in MOSFETs limit the degree to which the devices can be scaled down [1–3]. To overcome these limitations, it is required that the switching parameters of the devices be improved. In order to improve the device performance and minimize the short channel effects, various changes in the device design have been recommended such as strained silicon [4, 5], silicon on insulator (SOI), high k dielectric (HK), metal gate (MG), non-uniform doping. But, these were not sufficient and several other device structures like double-gate MOSFET, double-gate FINFET, TRI-gate FINFET were also proposed and investigated. FINFET is a promising substitute to MOSFET which is being recommended as the basis for future IC technologies because of its benefits such as low off-current, low standby power consumption, low sub-threshold swing

R. Koul (✉) · M. Yadav · R. Pandey
Delhi Technological University, New Delhi, India

value. However, as the devices continue to be scaled, more and more challenges prevent FINFET from achieving faster switching speed and lower power consumption [6]. Much improved performance characteristics can be achieved with the use of GAAFET device structures in which the insulating oxide and the gate electrode wrap around the channel material from all sides, thus achieving much better gate electrostatic control [7]. The analysis carried out in this work clearly shows that GAAFET gives best electrostatic control of the gate and minimizes the short channel effects along with excellent current switching ratio and low sub-threshold swing value. The analysis also shows that with the use of appropriate channel material and gate oxide material, the device performance can further be improved to a great extent. Further, analysis of the I_{ds} versus V_{ds} graphs clearly shows that GAAFET devices act as much better constant current sources in the saturation region than MOSFET and FINFET devices.

2 Device Structure and Description

All the devices under consideration consist of an $n+$ source and $n+$ drain region. The 2D/3D views of all the devices designed in this work are shown in Fig. 1. Here, Fig. 1a shows the 2D view of the MOSFET device, and the 3D view of the tri-gate FINFET designed in this work is displayed in Fig. 1b. The integration of tri-gate increases the gate controllability more than a double-gate FINFET, thereby allowing a higher drive current and a lower off-current along with a low sub-threshold swing value. Thick field oxide covering has been used for electrical isolation of the device. The 3D and cross-sectional view of the GAAFET devices modeled in this work is depicted in Fig. 1c. A two-channel device has been modeled which is completely wrapped up by the oxide material and the gate electrode, for achieving better results than a single-channel device. Field oxide covering has been used for electrical isolation of the device. The structural parameters of the modeled devices have been listed in Table 1.

3 Results and Discussions

The I_{ds} versus V_{gs} plots for MOSFET device using SiO_2 and HfO_2 as gate oxide for different channel materials are depicted in Figs. 2 and 3, respectively.

The I_{ds} versus V_{gs} plots for FINFET device using SiO_2 and HfO_2 as gate oxide for different channel materials are depicted in Figs. 4 and 5, respectively.

The plots obtained for the FINFET device using Si and SiGe as channel materials show nearly similar characteristics. Further, the values of drain current for the FINFET device using SiC_3C as channel material are again less in comparison to the others.

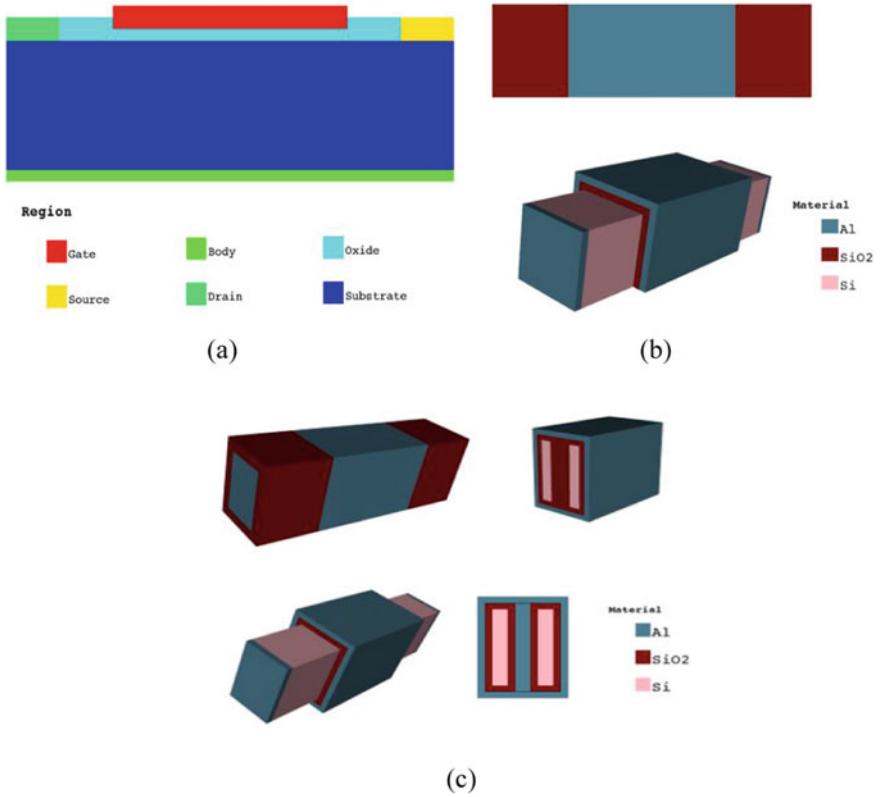


Fig. 1 a 2D view of MOSFET, b 3D view of FINFET, c 3D view of GAAFET

Table 1 Structural parameter of modeled device

Parameter	Device					
	MOSFET		FINFET		GAAFET	
Device dimensions	Gate length	22 nm	Gate length	22 nm	Gate Length	22 nm
			Width	14 nm	Width	16 nm
			Height	12 nm	Height	12 nm
Oxide thickness	1 nm		1 nm		1 nm	
Source and drain doping concentration	$1 \times 10^{20}/\text{cm}^3$		$1 \times 10^{20} /\text{cm}^3$		$1 \times 10^{20}/\text{cm}^3$	
Channel doping concentration	$1 \times 10^{16}/\text{cm}^3$		$1 \times 10^{16} /\text{cm}^3$		$1 \times 10^{16}/\text{cm}^3$	

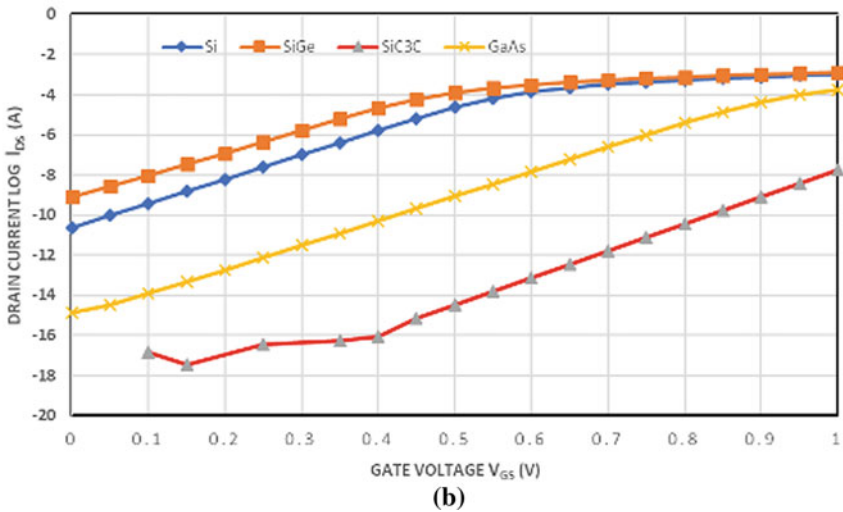
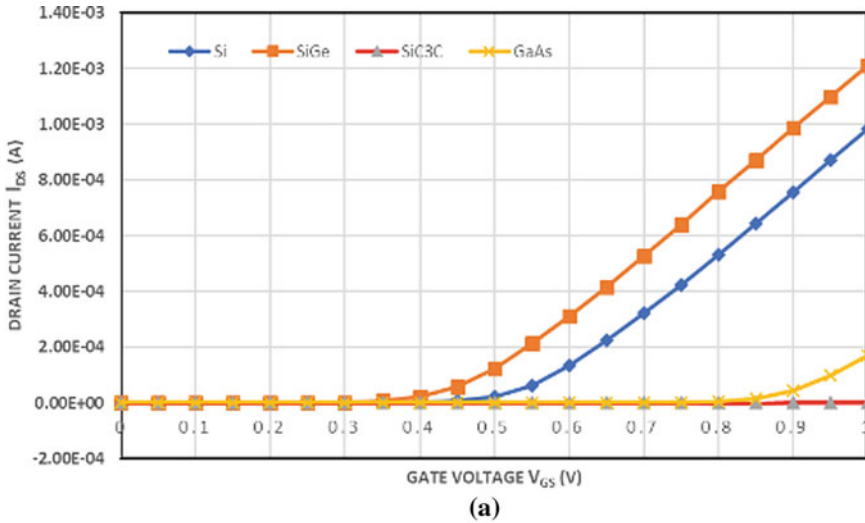


Fig. 2 a I_{ds} versus V_{gs} plot for MOSFET with SiO_2 as gate oxide and different channel materials. **b** $\log(I_{ds})$ versus V_{gs} plot for MOSFET with SiO_2 as gate oxide and different channel materials

The I_{ds} versus V_{gs} plots for GAAFET device using SiO_2 and HfO_2 as gate oxide for different channel materials are depicted in Figs. 6 and 7, respectively.

From these figures, it can be observed that the plots for the GAAFET device using Si and SiGe as channel materials also show nearly similar characteristics.

The comparison of the analog parameters, namely current switching ratio and sub-threshold swing, for the FINFET device using SiO_2 and HfO_2 as gate oxide has been depicted in Figs. 8 and 9, respectively. For comparative analysis, the on-current

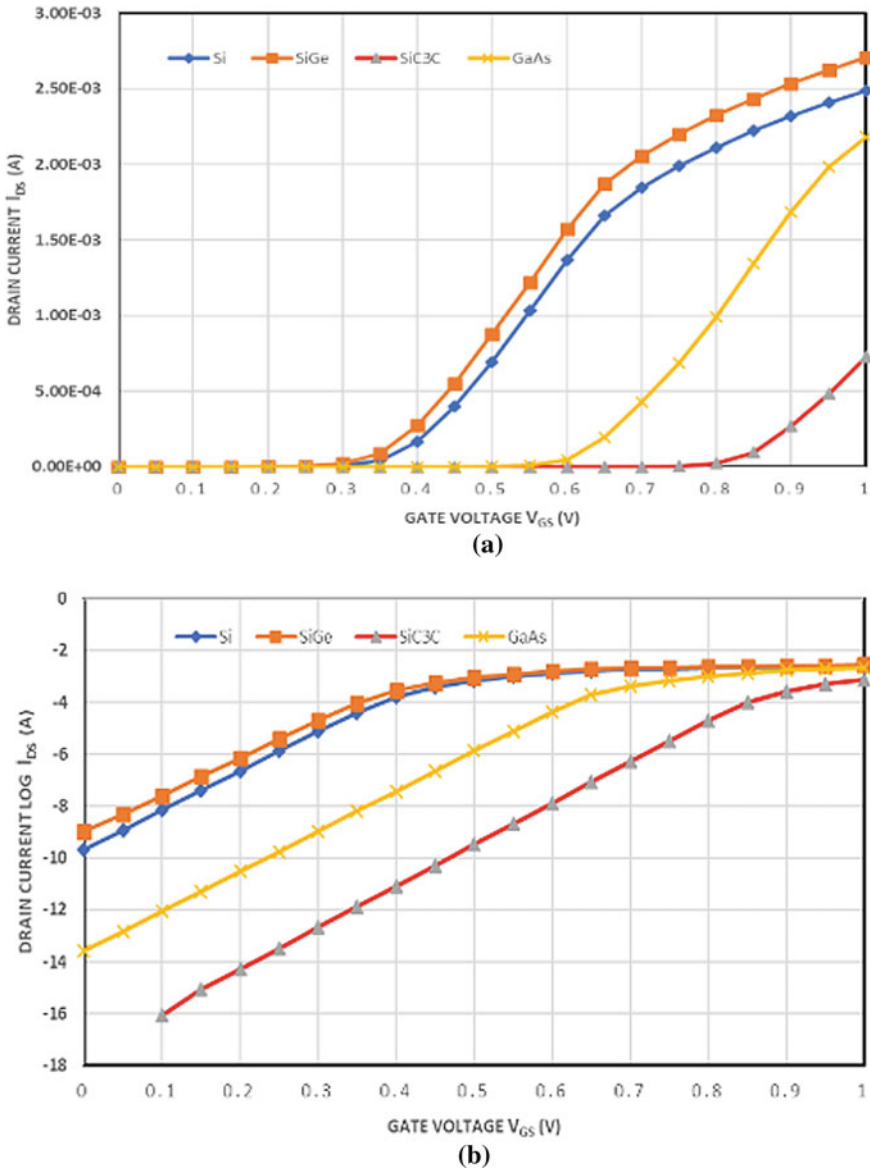


Fig. 3 a I_{ds} versus V_{gs} plot for MOSFET with HfO_2 as gate oxide and different channel materials. b $\log(I_{ds})$ versus V_{gs} plot for MOSFET with HfO_2 as gate oxide and different channel materials

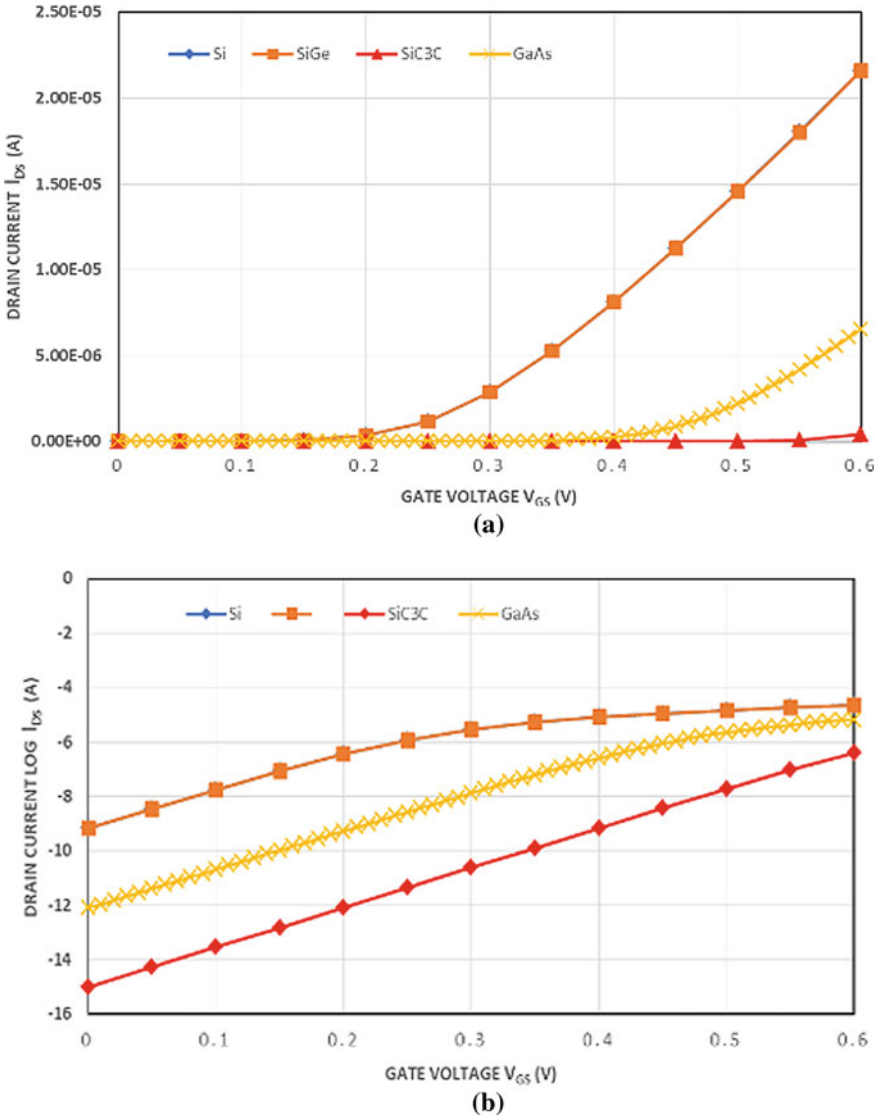
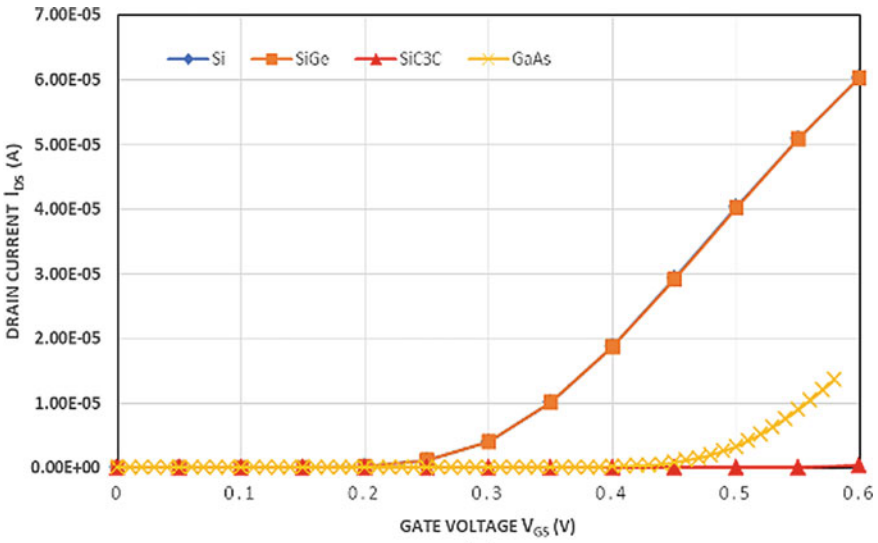


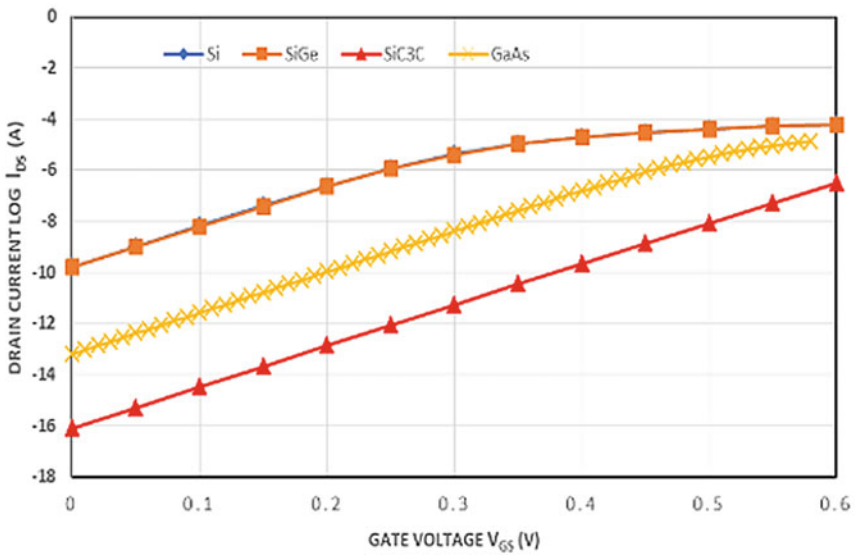
Fig. 4 a I_{ds} versus V_{gs} plot for FINFET with SiO_2 as gate oxide and different channel materials. b $\log(I_{ds})$ versus V_{gs} plot for FINFET with SiO_2 as gate oxide and different channel materials

of all the devices has been taken as the value of the drain current at $V_{gs} = 0.55$ V. For better visibility, log plots have been plotted (Figs. 10 and 11).

Detailed results of the simulations for all the devices with different gate oxides and substrate materials are summarized in Table 2.



(a)



(b)

Fig. 5 **a** I_{ds} versus V_{gs} plot for FINFET with HfO_2 as gate oxide and different channel materials. **b** $\log(I_{ds})$ versus V_{gs} plot for FINFET with HfO_2 as gate oxide and different channel materials

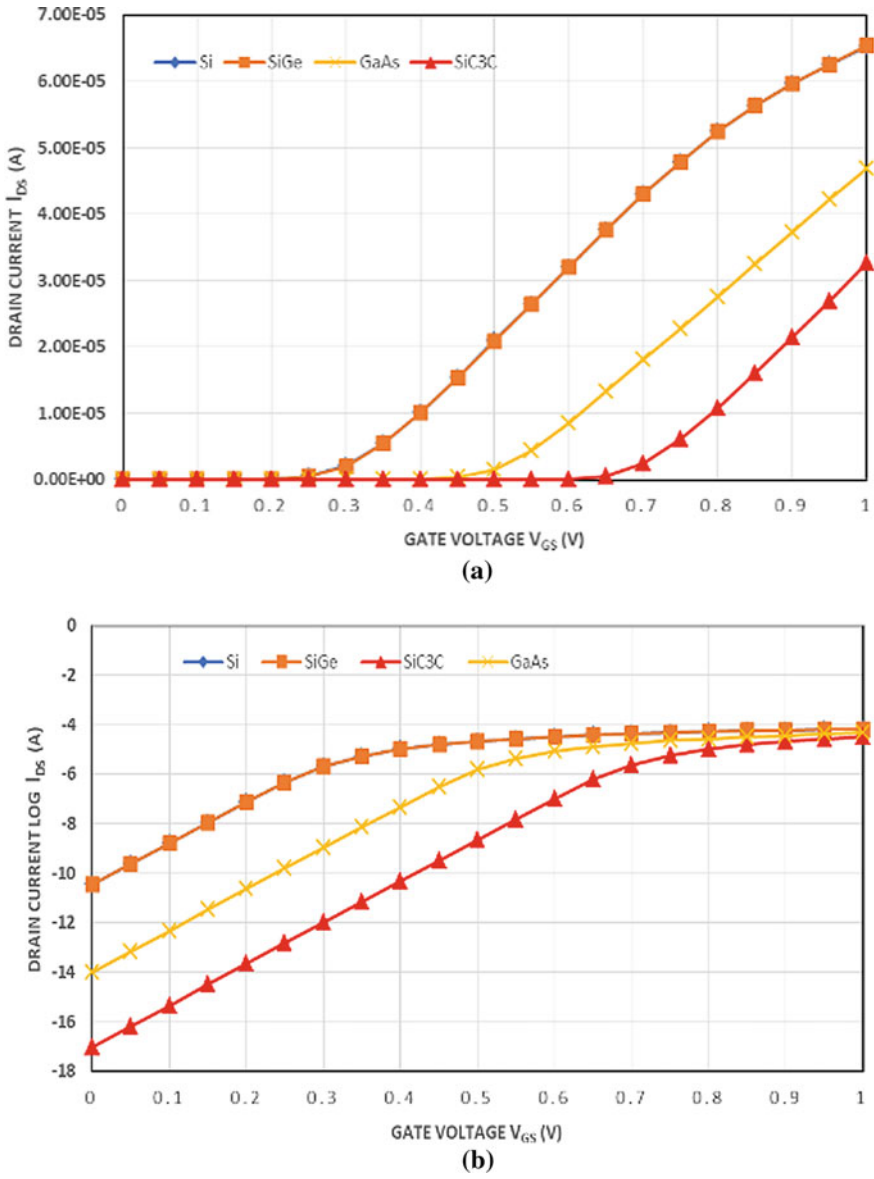


Fig. 6 **a** I_{ds} versus V_{gs} plot for GAAFET with SiO₂ as gate oxide and different channel materials. **b** Log(I_{ds}) versus V_{gs} plot for GAAFET with SiO₂ as gate oxide and different channel materials

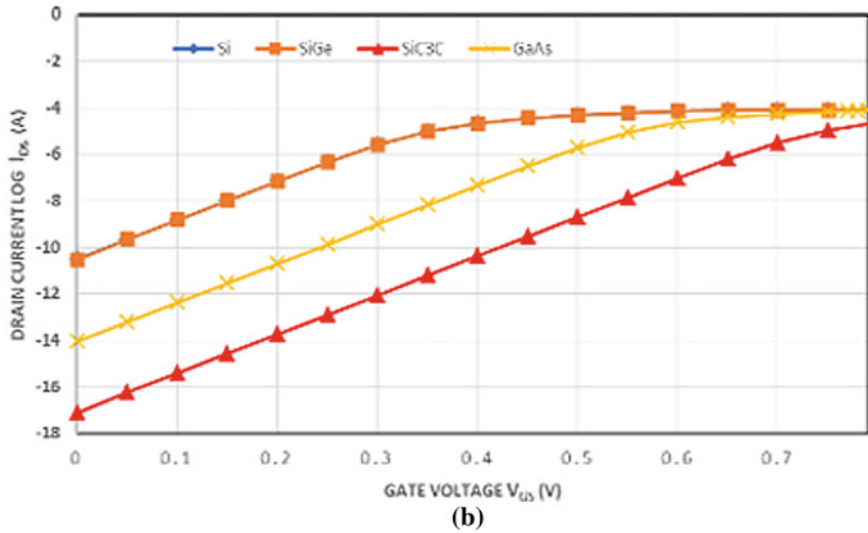
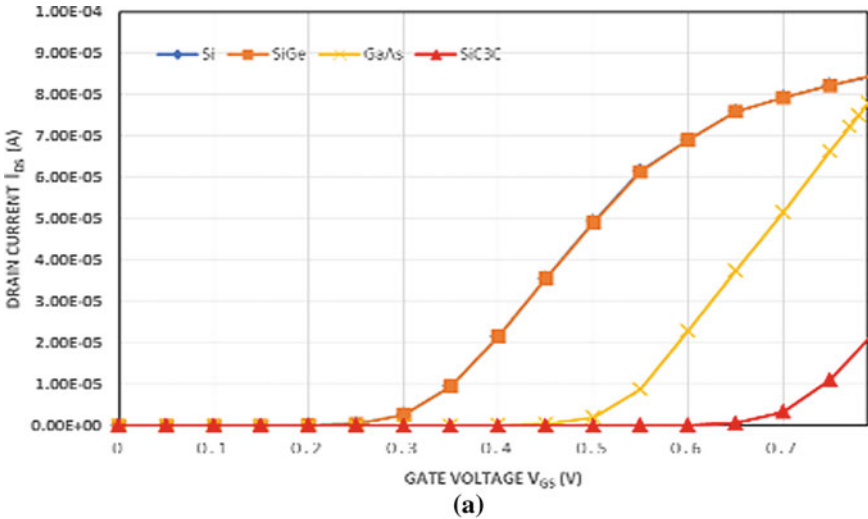


Fig. 7 a I_{ds} versus V_{gs} plot for GAAFET with HfO_2 as gate oxide and different channel materials. b $\log(I_{ds})$ versus V_{gs} plot for GAAFET with HfO_2 as gate oxide and different channel materials

It is observed that for the FINFET device that uses SiO_2 as the gate oxide, using Si and SiGe as channel materials results in the maximum on-current, whereas minimum off-current is obtained if SiC_3C is used as the channel material. Further use of SiC_3C also results in maximum current switching ratio (I_{on}/I_{off}) and minimum sub-threshold swing (SS) values.

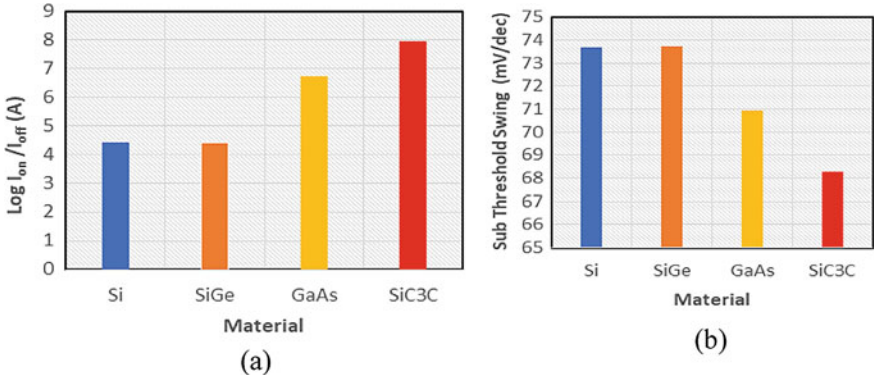


Fig. 8 **a** Comparison of current switching ratio for FINFET with SiO_2 as gate oxide. **b** Comparison of subthreshold swing for FINFET with SiO_2 as gate oxide

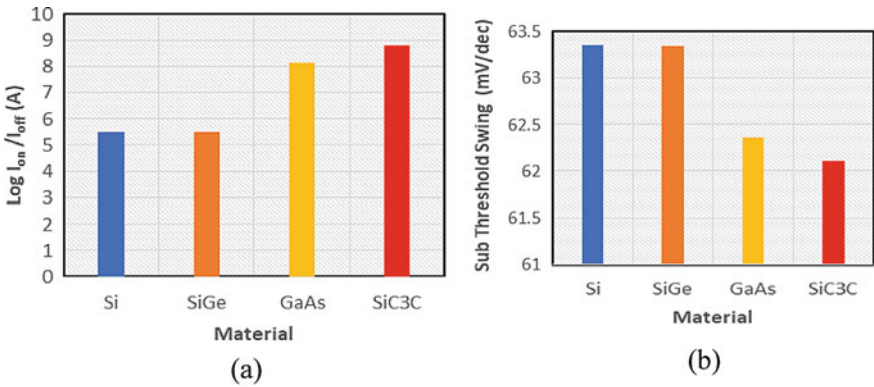


Fig. 9 **a** Comparison of current switching ratio for FINFET with HfO_2 as gate oxide. **b** Comparison of subthreshold swing for FINFET with HfO_2 as gate oxide

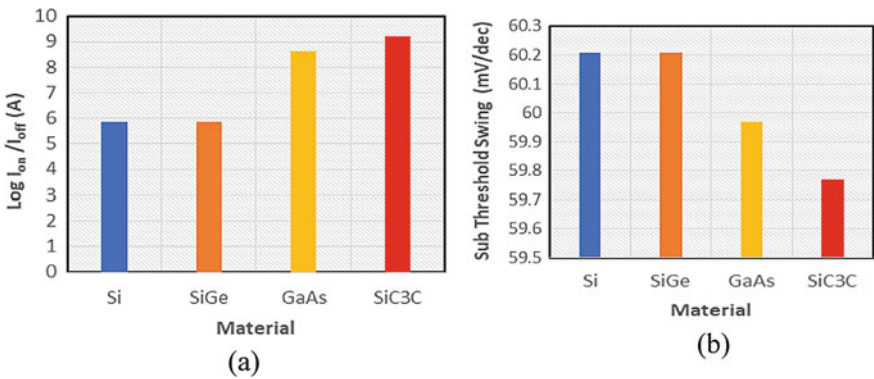


Fig. 10 **a** Comparison of current switching ratio for GAAFET with SiO_2 as gate oxide. **b** Comparison of subthreshold swing for GAAFET with SiO_2 as gate oxide

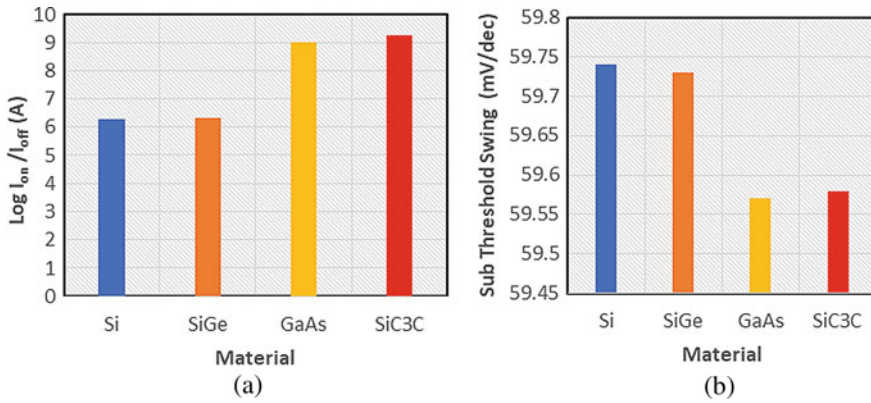


Fig. 11 a Comparison of current switching ratio for GAAFET with HfO_2 as gate oxide. **b** Comparison of subthreshold swing for GAAFET with HfO_2 as gate oxide

For the FINFET device that uses HfO_2 as the gate oxide, it is evident from Table 2 that using Si and SiGe, maximum on-current is achieved, whereas minimum off-current, maximum current switching ratio ($I_{\text{on}}/I_{\text{off}}$) and minimum sub-threshold swing values (SS) are obtained if SiC_3C is used as the channel material. Thus, it can be concluded that better performance is achieved with the use of HfO_2 rather than SiO_2 as the gate oxide as all the four analog parameters being analyzed show improvement.

For the GAAFET device that uses SiO_2 as the gate oxide, it is observed that although the trend is same as that for FINFET devices using SiO_2 as gate oxide the analog parameters being analyzed show significant improvement. The maximum on-current and current switching ratio achieved are greater than that of the FINFET device using SiO_2 as gate oxide, whereas the minimum off-current and sub-threshold swing are less. Further, the characteristics show improvement if HfO_2 is used as the gate oxide as evident from Table 2. A comparative study among the FINFET device using GaAs as channel material and HfO_2 as gate oxide used in this work and Ref. [8] shows that FINFET modeled in this work achieves a lower sub-threshold swing value and a higher current switching ratio (Table 3).

The plots for drain current (I_{ds} vs. V_{ds}) for MOSFET, FINFET and GAAFET devices using Si as channel material and SiO_2 as gate oxide have been depicted in Fig. 12. The simulations have been performed by keeping V_{gs} values fixed at 0.25 V, 0.27 V, 0.30 V, 0.32 V, respectively.

It is observed that the variation of current in the saturation region is least for the case of GAAFET which is the need for various applications. Thus, GAAFET can be used in the applications involving the use of constant current sources instead of MOSFET or FINFET devices.

Table 2 Comparison of MOSFET, FINFET and GAAFET devices

<i>DEVICE: MOSFET</i>					
GATE OXIDE		I_{on}	I_{off}	I_{on}/I_{off}	S.S
SiO ₂	Si	6.35E-05	2.33E-11	2.72E06	82.71
	SiGe	2.12E-04	7.60E-10	2.78E05	89.64
	GaAs	3.46E-09	1.31E-15	2.64E06	85.61
	SiC ₃ C	1.57E-14	3.01E-17	5.21E02	82.14
GATE OXIDE HfO ₂	Si	1.03E-03	2.12E-10	4.85E06	65.99
	SiGe	1.22E-03	1.03E-09	1.18E06	69.84
	GaAs	7.75E-06	2.56E-14	3.02E08	65.32
	SiC ₃ C	2.05E-09	2.68E-17	7.64E07	53.23
<i>DEVICE: FINFET</i>					
GATE OXIDE SiO ₂	Si	1.80E-05	6.91E-10	2.61E04	73.72
	SiGe	1.80E-05	6.86E-10	2.63E04	73.75
	GaAs	4.18E-06	7.89E-13	5.29E06	70.95
	SiC ₃ C	9.15E-08	9.75E-16	9.39E07	68.29
GATE OXIDE HfO ₂	Si	5.09E-05	1.65E-10	3.09E05	63.36
	SiGe	5.09E-05	1.61E-10	3.15E05	63.35
	GaAs	9.04E-06	6.43E-14	1.41E08	62.36
	SiC ₃ C	5.26E-08	8.01E-17	6.57E08	62.10
<i>DEVICE: GAAFET</i>					
GATE OXIDE SiO ₂	Si	2.64E-05	3.65E-11	7.24E05	60.21
	SiGe	2.64E-05	3.58E-11	7.37E05	60.21
	GaAs	4.41E-06	1.05E-14	4.19E08	59.97
	SiC ₃ C	1.52E-08	9.48E-18	1.60E09	59.77
GATE OXIDE HfO ₂	Si	6.15E-05	3.09E-11	1.99E06	59.73
	SiGe	6.14E-05	3.02E-11	2.03E06	59.73
	GaAs	8.75E-06	8.74E-15	1.00E09	59.56
	SiC ₃ C	1.38E-08	7.82E-18	1.76E09	59.57

Table 3 Comparison of FINFET device with Ref. [8]

Electrical parameters	FINFET using GaAs as channel material and HfO ₂ as gate oxide in this work	Reference [8]
1. Sub-threshold swing (SS)	62.36	63.21
2. Current switching ratio (I_{on}/I_{off})	1.41E08	2.75E07

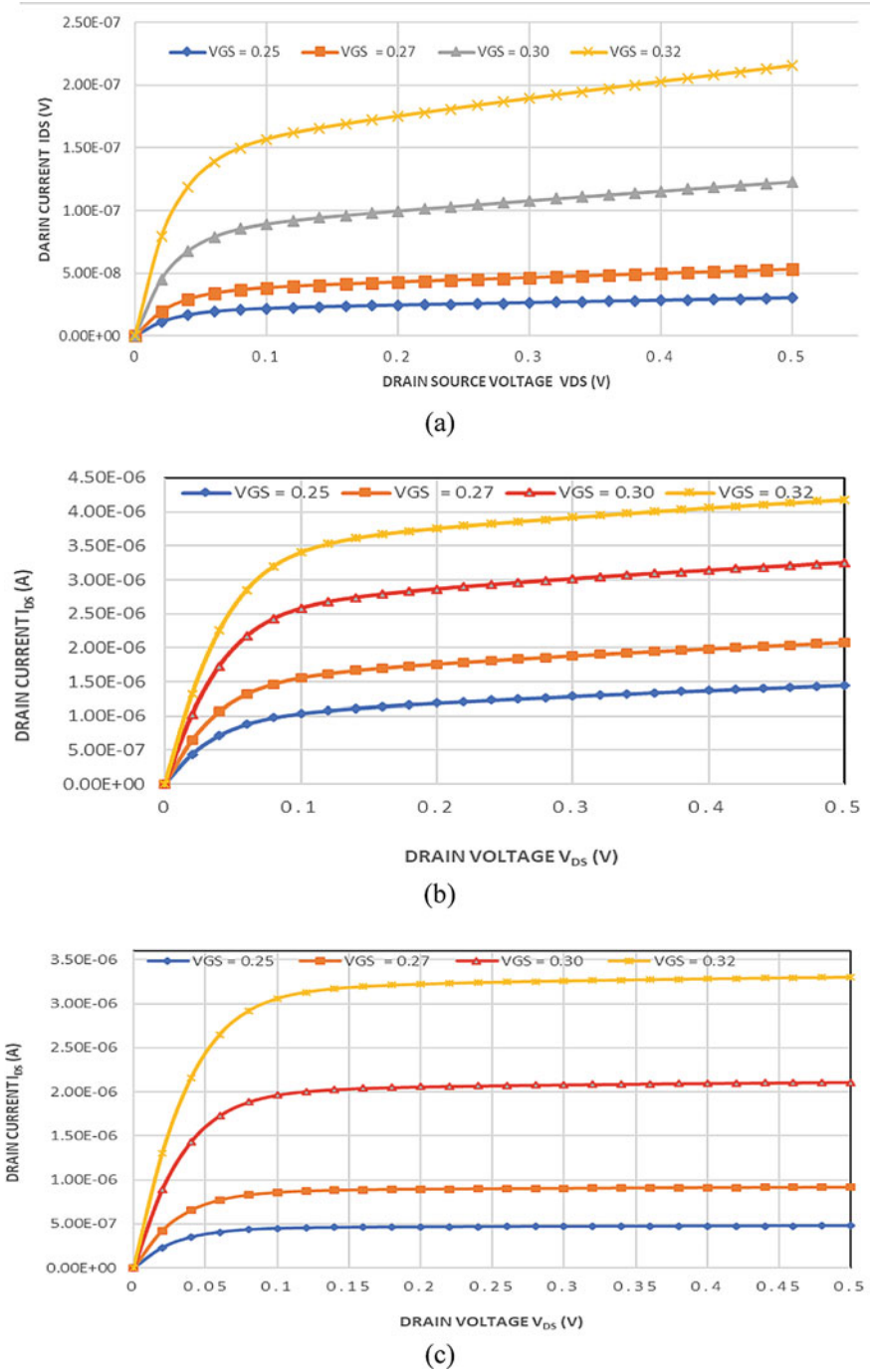


Fig. 12 I_{ds} versus V_{ds} plot for a MOSFET, b FINFET, c GAAFET

4 Conclusion

The influence of varying channel material and gate oxide material for MOSFET, FINFET and GAAFET devices on off-current, on-current, current switching ratio and sub-threshold swing has been explained. It is found that use of HfO_2 as gate oxide improves the performance of the device and use of appropriate channel material according to the application can help to achieve optimum characteristics. Also, a comparison of MOSFET, FINFET and GAAFET devices as current sources has been presented from which it is concluded that GAAFET devices can act as best constant current sources among these.

References

1. P.-F. Wang, K. Hilsenbeck, T. Nirschl, M. Oswald, C. Stepper, M. Weis et al., Complementary tunneling transistor for low power application. *Solid-State Electron.* **48**, 2281–2286 (2004)
2. J. Madan, K. Karwal, R. Chaujar, Performance analysis of heterojunction DMDG-TFET with different source materials for analog application, in *2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, 2018, pp. 1474–1478
3. A. Sarkar, S. Halim, A. Ghosh, S. Sarkar, Implementation of PMN-PT/Ni based NOR Gate with biaxial anisotropy off ultra low energy dissipation. *J. Nanoelectron. Optoelectron.* **12**, 1–6 (2017)
4. A. Khakifirooz, K. Cheng, T. Nagumo, N. Loubet, T. Adam, A. Reznicek, J. Kuss, D. Shahrjerdi, R. Sreenivasan, S. Ponoht, H. He, Strain engineered extremely thin SOI (ETSOI) for high-performance CMOS, in *2012 Symposium on VLSI Technology (VLSIT)* (IEEE, 2012), pp. 117–118
5. P. Banerjee, P. Saha, D.K. Dash, A. Ghosh, S.K. Sarkar, Analytical modeling and performance analysis of graded channel strained dual-material double gate MOSFET, in *4th International Conference on Computing Communication and Automation 2018 (ICCCA)*
6. Y. Zhuo, et al., Statistical variability analysis in vertically stacked gate all around FETs at 7 nm technology, in *2018 14th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, Qingdao, 2018, pp. 1–3
7. M.S. Yeh, Y.J. Lee, M.F. Hung, K.C. Liu, Y.C. Wu, High-performance gate-all-around poly-Si thin-film transistors by microwave annealing with NH_3 plasma passivation. *IEEE Trans. Nanotechnol.* **12**, 636
8. R. Saha, B. Bhowmick, S. Baishya, Effects of temperature on electrical parameters in GaAs SOI FinFET and application as digital inverter, in *2017 Devices for Integrated Circuit (DevIC)*, Kalyani, 2017, pp. 462–466

VLSI Design of a Split Parallel Two-Dimensional HEVC Transform



Ainy Haziyah Awab, Ab Al-Hadi Ab Rahman, Izam Kamisian,
and Mohd Shahrizal Rusli

Abstract This paper proposes a highly parallel two-dimensional (2D) HEVC transform hardware architecture, implemented in 32-nm VLSI technology. The design allows very high-resolution and frame-rate video coding by way of a very fast HEVC transform operations. It is based on a split architecture, where the individual transform type and size is separated into its own core, therefore enables pixel-level parallelism in the 2D parallel and folded structures. This work also implements the full specification of the HEVC transform for both the DCT and DST transforms, with performance, power, and area analyses for the two structures. Results show very significant speed up over existing unified architectures, with only a relatively modest increase in total gate count. The design is suitable for applications that require very high video resolution and frame rate.

Keywords HEVC transform · DCT · DST · VLSI

1 Introduction

High-efficiency video coding (HEVC) is the current video compression standard, developed by the Joint Collaborative Team on Video Coding (JCT-VC), a collaboration between the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) standardization organizations [1]. It is widely used today to encode and decode videos for playback in devices such as

A. H. Awab (✉) · A. A.-H. A. Rahman · I. Kamisian · M. S. Rusli
Universiti Teknologi Malaysia, 81310 Johor Bahru, Malaysia
e-mail: ahaziyah2@graduate.utm.my

A. A.-H. A. Rahman
e-mail: hadi@utm.my

I. Kamisian
e-mail: e-izam@utm.my

M. S. Rusli
e-mail: shahrizal@utm.my

Table 1 Number of arithmetic and logical shift operations in an N -point butterfly DCT/DST

	Multiplication	Addition	Shifter
4-Point DCT/DST	8	12	8
8-Point DCT	24	36	16
16-Point DCT	88	100	32
32-Point DCT	344	404	64

computers and smartphones. One of the core components in HEVC is its transform, known to be very compute-intensive and power-hungry. For usage of HEVC in mobile devices and embedded systems, it is crucial that the operation should be efficient in terms of power, while able to support high frame-rate and high-resolution videos [2].

The HEVC transform, as defined in the standard, specifies two-dimensional (2D) discrete cosine transform (DCT) of size $N = \{4, 8, 16, 32\}$ and discrete sine transform (DST) of size $N = 4$ [3]. The 2D operation is performed by multiplying each row of the input image block of size N by an $N \times N$ DCT or $N \times N$ DST matrix for the first dimension, followed by each column for the second dimension. The estimated number of arithmetic and logic operations for each row or column of N -point transform type and size for a butterfly decomposition structure [4] is given in Table 1. The complete 2D transform of image block size of $N \times N$ thus requires $2 \times N \times N$ -point operations.

Due to the high computational requirements of the HEVC transform, several works have been reported in the design of a custom hardware for the operation. The goal is mainly to leverage the compute-intensive transform operations to dedicated hardware (e.g., FPGA or ASIC), while maintaining other serial-based HEVC operations in software.

Some of the important works include an integer-based multiplierless constant multiplication (MCM) implementation [5], and the Walsh–Hadamard transform (WHT) [6]. In [5], the authors use the MCM approach for 4-point and 8-point DCT, while normal multiplication was used for 16-point and 32-point DCT. In [6], the authors use the Given rotations the 2D N -point DCT, which are cascaded with the DCT matrix. Both of these works utilize two commonly used structures: the parallel and folded structures. The parallel structure utilizes two separate cores for the two dimensions, while the folded uses only a single transform core. These works obtain good power and energy figures, but may be limited in terms of performance since some data parallelism is constrained as the unified architecture is used.

Apart from this, other reported works include a pipelined implementation of the DCT/DST transform [7–9] with the objective of enhancing the performance. Another interesting work is given in [10] where the DCT is optimized for the most likely block size in the video stream. The work in [11] utilizes the carry save adder-based multiplier unit for improved performance, while the work in [12] explores the design space of the HEVC transform at high level using the methodology described in [13]. The objective in this case is to find the best trade-off between performance, power, and area.

The contribution of this paper is on the hardware design of a highly parallel HEVC transform using the proposed top-level split architecture. Furthermore, the parallel and folded structures are also extended and integrated in the 2D transform core, and the top-level analyzed for performance, power, and area [14–16]. This work also implements the complete design of the HEVC transform with all possible types and sizes of the transform, with comparison of results from the works in [5, 6].

2 Proposed Architectures

2.1 Top-Level Split Architecture

The top-level block diagram of the proposed split architecture is given in Fig. 1. Essentially, each 2D DCT transform unit of size $N = \{4, 8, 16, 32\}$ and DST of size $N = 4$ is separated into its own core. Compared to the unified architectures, this may seem to result in large area and possibly consumes more power. However, there are two advantages that can be expected from this architecture:

1. Two blocks of distinct sizes can be processed at the same time, whereas in the unified architecture, a new block can only be processed after the completion of the current block.
2. As each transform core of size N is designed separately, it is now possible to design with pixel-level parallelism, as opposed to the unified architecture with row-level parallelism for each of the 2D transform unit.

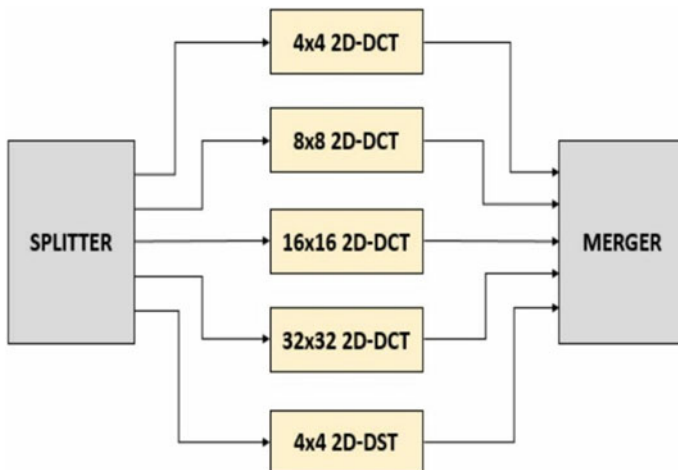


Fig. 1 The proposed split architecture for hardware implementation of the full HEVC 2D transform

The splitter takes the specification of the block (i.e., the size N and transform type) and routes the block accordingly to the relevant transform core. If a new block is available, it compares the size and type of the new block with the current block. If they are not equal, then the new block can be sent to the relevant transform core for concurrent processing. In the case that they are equal, then it has to wait until the current block is completed before routing the next block to the same transform core.

The transformed blocks are then sent to the merger, which essentially stores the results of the 2D DCT computations temporarily, before being routed to the next stage of the HEVC encoder/decoder.

2.2 2D Transform Architecture

In this subsection, we discuss the implementation of the individual 2D transform unit. The two most commonly used structures as discussed in Sect. 1 are implemented, namely the parallel and folded structures. However, since the split architecture is utilized, it is now possible to obtain maximum parallelism (i.e., pixel level) for the 2D transform. Algorithm 1 shows the operation for 2D transform where rows are first transformed to obtain 1D transform, followed by the transposition operation, and finally a second round of row transform to obtain the 2D transform.

Algorithm 1: 2D DCT/DST Transform of image block size $N \times N$

```

Input:  $XN$ : Input image block of size  $N \times N$ ,  $MN$ : DCT/DST matrix of
         size  $N \times N$ 
Output:  $YN_{\_2D}$ : 2D transformed image block of size  $N \times N$ 
for each row  $r = 1..N$  in  $XN$  do
  |  $YN\_1Dr = XNr * MN$ 
end
 $YN\_1DT = \text{Transpose}(YN\_1D)$ 
for each row  $r = 1..N$  in  $YN\_1DT$  do
  |  $YN\_2Dr = YN\_1DTr * MN$ 
end
return  $YN\_2D$ 

```

In the proposed split architecture, size N is explicitly defined for each core. Therefore, the two *for* loops in the algorithm can be unrolled, followed by a direct connection between first and second row transformations. Therefore, the transposition buffer can be removed from the algorithm which allows full parallelism for each pixel in both transform dimensions.

The hardware architectures for the 2D transform unit are given in Figs. 2 and 3 for parallel and folded structures, respectively. In the parallel structure, the operation is purely combinational logic with mainly adders and multipliers, without a transposition buffer. In the folded structure, however, a transposition buffer is required

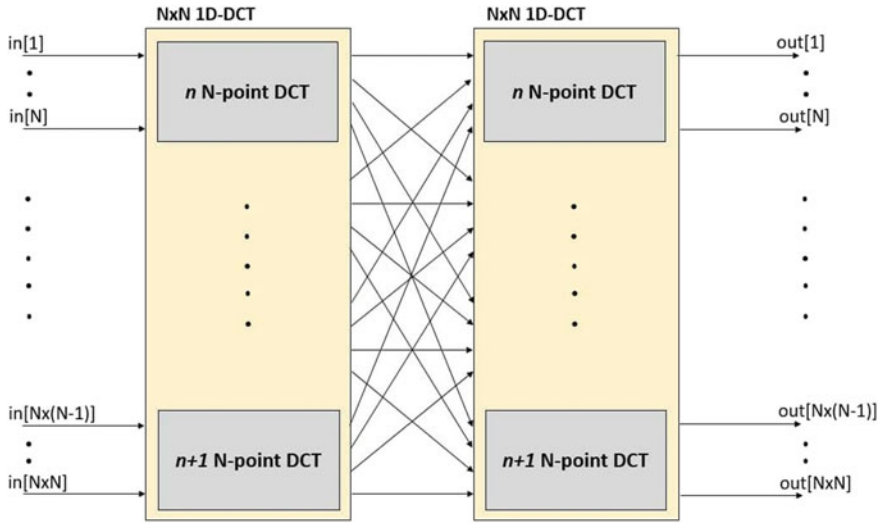


Fig. 2 Parallel hardware architecture for 2D DCT of size N

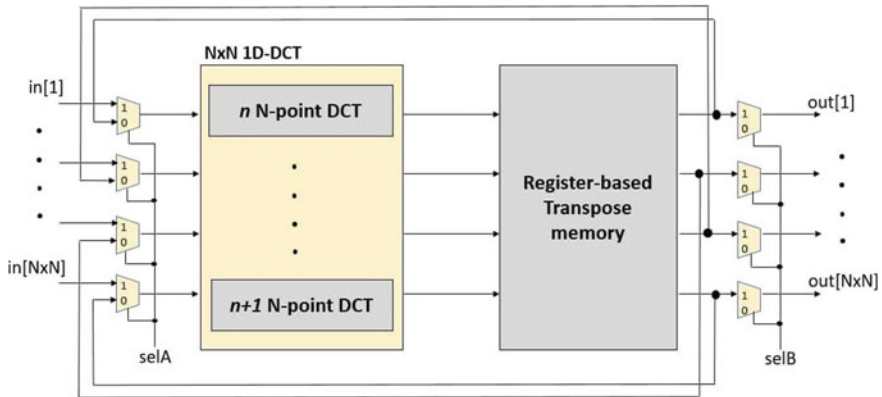


Fig. 3 Folded hardware architecture for 2D DCT of size N

as only a single 1D transform unit is allowed. In contrast to previous works, the 1D transform unit loop is fully unrolled for full $N \times N$ pixel parallelism.

3 Experimental Results

Each of the 2D DCT of size $N = \{4, 8, 16, 32\}$ and DST of size $N = 4$ has been implemented using the proposed split architecture for both parallel and folded structures. Verilog HDL has been used for RTL design, while EDA tools Synopsys DC and ICC have been used for front-end synthesis and back-end layout implementation using SAED 32nm process technology, respectively. The design physical layout has been extracted for parasitic and analyzed for timing and power. Other performance parameters analyzed include design gate count and throughput.

The graph in Fig. 4 presents the maximum operating frequency (fmax) for each DCT and DST for both parallel and folded structures. The average fmax is 385 MHz and 250 MHz, respectively, for folded and parallel structures. The fmax is found to be higher in the folded structure as expected, as the 2D transform data path is partitioned into two parts. In can be seen that the parallel structure has lower fmax by around 35% as compared to the folded structure.

For the parallel structure, the DCT and DST designs have latency of 1 clock cycle, while latency is 2 clock cycles for the folded structure. For practical purposes, a serializer–deserializer (SerDes) has also been integrated at the input and output of each 2D transform core. Using these, together with the fmax and the number of pixels for DCT and DST of size N , the throughput can be calculated. The result is shown in the graph of Fig. 5.

Figure 6 shows the total power consumption (static and dynamic) for each DCT and DST for both parallel and folded structures. Overall, it can be seen that the folded structures consume slightly more power compared to the parallel structures as the folded structure has significant number of registers. It is also interesting to see that the power is increased exponentially going from smaller to bigger size as the number of pixels being processed in parallel is also increased exponentially. This is

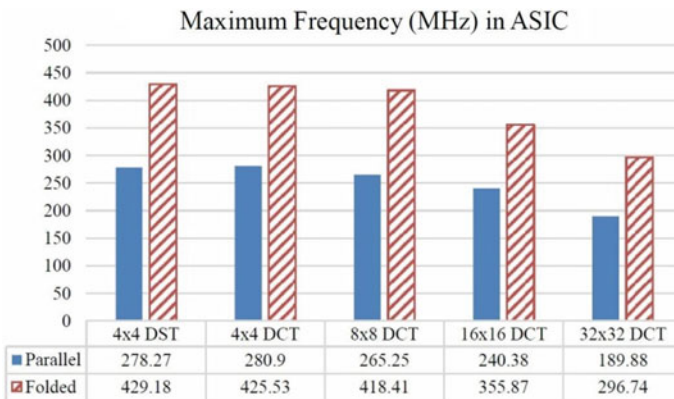


Fig. 4 Maximum operating frequency for each DCT and DST for both parallel and folded structures

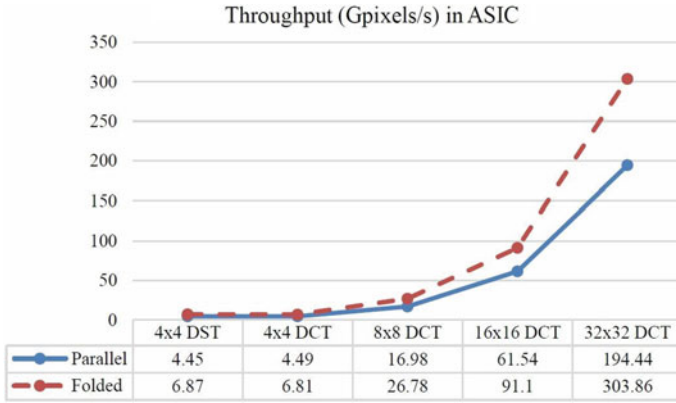


Fig. 5 Throughput for each DCT and DST for both parallel and folded structures

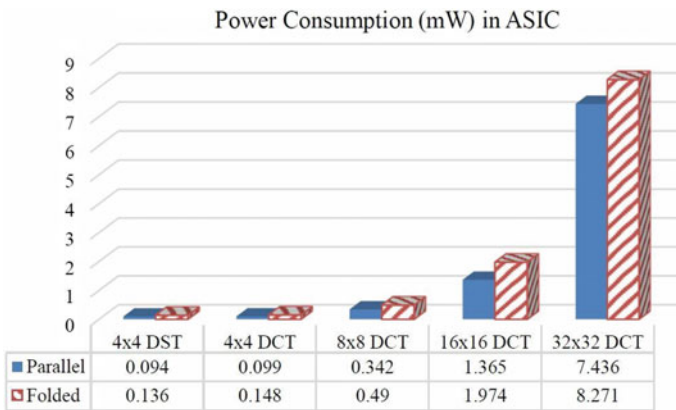


Fig. 6 Total power consumption for each DCT and DST for both parallel and folded structures

an interesting observation that while the folded structure results in higher throughput, power is more optimum in the parallel structure.

The gate count for each design is shown in Fig. 7. As expected, larger transform size utilizes higher number of cells, as shown with almost 1 million logic gates for the parallel 32×32 DCT. Similar to power consumption, the gate count can be seen to increase exponentially as the transform size is increased from $N = 4$ to $N = 32$.

The proposed work has been compared with similar works in the literature. As given in Table 2, the fmax is around the same values for all designs. As for throughput, we have obtained up to 65 times and 233 times higher compared to previous works for parallel and folder structures, respectively, mainly due to having highest level of parallelism using the split architecture. Because of this, the gate count in our case is also higher by up to 4.3 times and 3.8 times for the parallel and folded structures, respectively. In terms of power, our designs show considerably less power consumed,

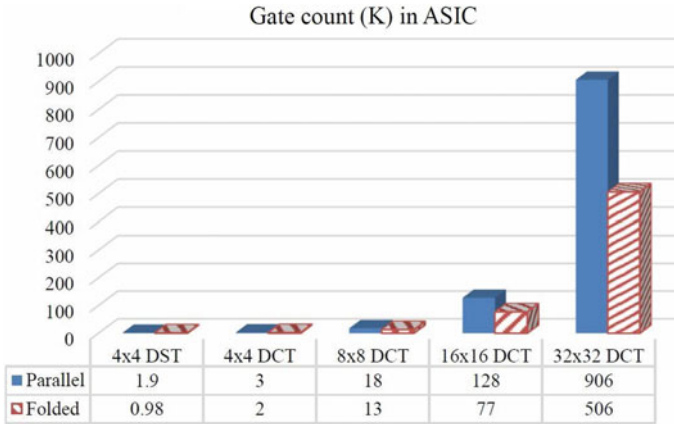


Fig. 7 Gate count for each DCT and DST for both parallel and folded structures

Table 2 Results comparison with similar works in literature

	[5] Parallel	[5] Folded	[6] Parallel	[6] Folded	Split Parallel	Split Folded
Technology (nm)	90	90	90	90	32	32
Fmax (MHz)	187	187	250	250	190	297
Thr. (Gpix/s)	5.984	2.992	3.212	1.302	194.4	303.9
Power (mW)	67.57	40.04	51.72	28.98	9.242	10.88
Gate count (K)	347	208	243	157	1055	598

possibly due to a more advanced process technology, by around 82% and 31% for parallel and folded structures, respectively.

All of these designs (prior works and ours) are able to support real-time video at 8 K resolution. The advantage of our proposed solution is the very high throughput (with acceptable increase in gate count), which allows the reduction in operating frequency and reduced power consumption for a given throughput requirement. Furthermore, the proposed design could also support HEVC video coding that requires higher video resolution and frame rate than current state of the art.

4 Conclusion

In this work, a parallel split architecture for VLSI implementation has been proposed and presented that allows pixel-level parallelism for each 2D transform core in the HEVC transform. The full HEVC transform for all sizes for DCT and DST has been designed, implemented, and analyzed for performance, power, and area.

Results show best operating frequency for the folded structure as expected, although with higher latency as compared to the parallel structure. Including the SerDes, the folded structure is found to perform slightly better in terms of throughput. In terms of power, however, the parallel structure performs better, mainly due to the absence of registers in the design. In terms of area, the parallel structure shows higher gate count due to a higher utilization of arithmetic units. As compared with previous works using unified architectures, the proposed split architecture results in significantly higher throughput with only a modest increase in gate count.

As for future work, it would be interesting to analyze the performance using real video stream benchmarks where the 2D transform core can be run in parallel. This would also require the integration of the proposed hardware HEVC transform with other components in HEVC codec either in software or hardware.

References

1. G.J. Sullivan, J.-R. Ohm, W.-J. Han, T. Wiegand, Overview of the high efficiency video coding (HEVC) standard. *IEEE Trans. Circuits Syst. Video Technol.* **22**(12), 16491668 (2012)
2. V. Sze, M. Budagavi, G.J. Sullivan, High efficiency video coding (HEVC), in *Integrated Circuit and Systems, Algorithms and Architectures* (Springer, Berlin, 2014), p. 1375
3. M. Budagavi, A. Fuldseth, G. Bjntegaard, V. Sze, M. Sadafale, Core transform design in the high efficiency video coding (HEVC) standard. *IEEE J. Sel. Top. Signal Process.* **7**(6), 10291041 (2013)
4. T. Ma, C. Liu, Y. Fan, X. Zeng, A fast 8×8 IDCT algorithm for HEVC, in *2013 IEEE International Conference on ASIC* (2013), pp. 1–4
5. P.K. Meher, S.Y. Park, B.K. Mohanty, K.S. Lim, C. Yeo, Efficient integer DCT architectures for HEVC. *IEEE Trans. Circuits Syst. Video Technol.* **24**(1), 168178 (2014)
6. M. Maserà, M. Martina, G. Maserà, Area efficient DST architectures for HEVC. 101–104 (2017). <https://doi.org/10.1109/PRIME.2017.7974117>
7. J. Zhu, Z. Liu, D. Wang, Fully pipelined DCT/IDCT/Hadamard unified transform architecture for HEVC codec, in *Proceedings IEEE International Symposium Circuits and Systems (ISCAS)*, May 2013 (2013), p. 677680
8. J. Nan, N. Yu, W. Lu, D. Wang, A DST hardware structure of HEVC, in *2nd International Conference on Information Science and Control Engineering*, Shanghai (2015), pp. 546–549
9. A. Sarkar, S. Halim, A. Ghosh, S. Sarkar, Implementation of PMN-PT/Ni based NOR Gate with biaxial anisotropy off ultra low energy dissipation. *J. Nanoelectron. Optoelectron.* **12**, 1–6 (2017)
10. A. Sen, A. Ghosh, Performance study of PMNPT/Ni-based multiferroic nanomagnet for straintronics device applications, in *Advances in Intelligent Systems and Computing*, ed. by L.C. Jain, M. Virvou, V. Piuri, V.E. Balas (Springer, Berlin, 2019), pp. 15–24. ISSN: 2194-5357
11. P. Banerjee, P. Saha, D.K. Dash, A. Ghosh, S.K. Sarkar, Analytical modeling and performance analysis of graded channel strained dual-material double gate MOSFET, in *4th International Conference on Computing Communication and Automation 2018 (ICCCA)*
12. N.M. Zabidi, A.A.H. Ab-Rahman, VLSI design of a fast pipelined 8×8 discrete cosine transform. *Int. J. Electr. Comput. Eng.* 1430–1440 (2017)
13. N.C. Vayalil, J. Hadrill, Y. Kong, An efficient ASIC design of variable-length discrete cosine transform for HEVC, in *2016 European Modelling Symposium (EMS)*, Pisa (2016), pp. 229–233
14. M. Mohamed Asan Basiri, N.M. Sk, High performance integer DCT architectures for HEVC, in *30th International Conference on VLSI Design and 2017 16th International Conference on Embedded Systems (VLSID)*, Hyderabad (2017), pp. 121–126

15. K.Z. Yion, A.A.H. Ab-Rahman, Exploring the design space of HEVC inverse transforms with dataflow programming. *Indones. J. Electr. Eng. Comput. Sci.* **6**, 104 (2017). <https://doi.org/10.11591/ijeecs.v6.i1.pp104-109>
16. H. Amer, A.A.H. Ab-Rahman, I. Amer, C. Lucarz, M. Mattavelli, Methodology and technique to improve throughput of FPGA-based cal dataflow programs: case study of the RVC MPEG-4 SP intra decoder, in *2011 IEEE Workshop on Signal Processing Systems (SiPS)* (2011), pp. 186–191

Convergence of Novel Iterative Learning Control Methods for a Class of Linear Discrete-Time Switched Systems



D. R. Sahu and Nitish Kumar Singh

Abstract In this paper, we study the convergence of different iterative learning control methods implemented to a class of linear discrete-time switched systems coupled with an arbitrary switched rules under the hypothesis that switched system operates repetitively over a finite time interval. The convergence is discussed by utilizing the super-vector approach when the system is noise-free. Here, we introduce Mann-ILC method and normal S-ILC method of switched systems and establish some theorems related to the convergence of the corresponding reference trajectory on the time interval. We present a numerical experiment to demonstrate the convergence behaviour of theoretical analysis.

Keywords ILC · switched system · Super-vector approach · Mann iteration method · Normal S-iteration method

1 Introduction

In modern era, switched systems have acquired considerable attention due to their broad applications in the control of chemical systems, automotive industry, process control, power systems, networked control systems, engineering [1, 2] and many others [3–5]. Currently study of switching system mainly focussed on stability, controllability and observability [6–8]. The switched system is a collection of subsystems of a dynamical system and a logical rule which arrange switching between these subsystems. Precisely, these subsystems are generally characterized by difference equations or indexed differentials. The logical rule that organizes switching between these subsystems achieves switching signal, which is generally described as classes of piecewise constant maps,

$$\sigma : \mathbb{R}^+ \longrightarrow \mathbb{I},$$

D. R. Sahu · N. K. Singh (✉)
Department of Mathematics, Institute of Science,
Banaras Hindu University, Varanasi 221005, India

or, sequences

$$\sigma : \mathbb{Z}^+ \longrightarrow \mathbb{I}.$$

By wishing a switching signal to be piecewise constant, we focus that on any finite interval of \mathbb{R}^+ (positive real numbers) the switching signal σ has a finite number of discontinuities [6]. Although the study of the switching system is flourishing, generally the dynamics of the system (plant model) in engineering practice is unknown.

It is very demanding to design an effective model for the tracking performance of the switched system. Positively, there exists an effective control method called iterative learning control (ILC), which is a repeated control method that improves the tracking performance of the closed control system by using a learning capacity over repeated test trials as shown in Fig. 1. The mechanism associate designs an inverse model and estimates the tracking error of the earlier iteration and then calculates a corrective response from the tracking error and earlier input signal to obtain a revised input signal. The revised input signal is designed to improve the tracking performance of the next test trial, subsequently the tracking an error may be designed again and the process is repeated. The reference trajectory can be tracked under satisfactory conditions for convergence for associated method. One of the superiorities of ILC is that it requires less knowledge of the focussing system. By taking account the above property of ILC, it can be used in the switched system for tracking the desired result [9]. Arimoto et al. [10] were first to proposed the original form of ILC method in 1984, for more detail see [11, 12].

In [3], convergence property for a P-type ILC method has been analysed by super-vector approach for a class of linear discrete-time switched systems.

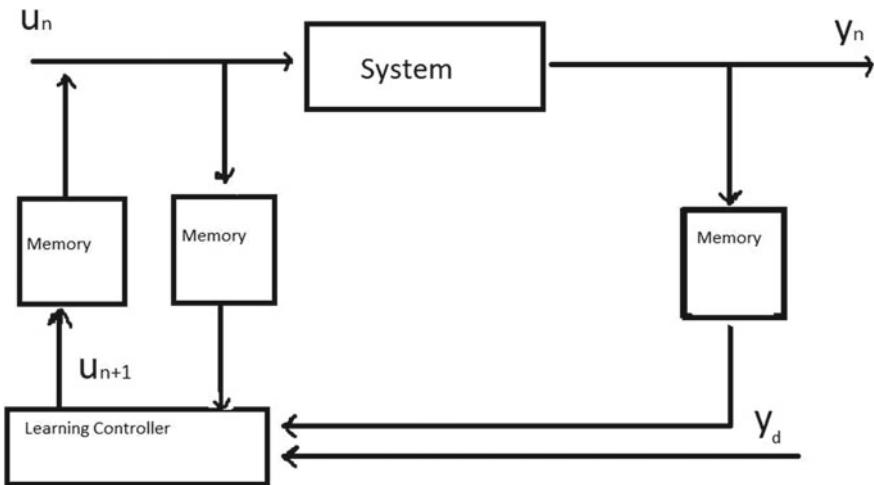


Fig. 1 ILC system

On the other side, in the area of nonlinear analysis, the fixed-point theory is an important and very interesting field. Many problems in mathematical physics and operational research can be modelled as a fixed point problem. In literature, many iterative techniques have been introduced and study for a fixed point of different systems.

Let C be a non-empty convex subset of a linear space X and $\mathcal{T} : C \rightarrow C$ is a mapping. For any initial point $u_0 \in C$, then Picard iteration (1.1) given by Picard [13], and Mann iteration (1.2) introduced by Mann [14], is, respectively, given by

$$u_{n+1} = \mathcal{T}(u_n), \quad \text{for all } n \in \mathbb{N}, \tag{1.1}$$

$$u_{n+1} = (1 - \alpha_n)u_n + \alpha_n\mathcal{T}u_n, \quad \text{for all } n \in \mathbb{N}, \tag{1.2}$$

where sequence $\{\alpha_n\}$ in $(0, 1)$ is iteration parameter. But it is the rate of convergence is slow, especially for the large-scale problem. To get rid of this problem, many authors introduced many iterations such as Ishikawa [15] and Krasnoselskii. In 2007, Agrawal, O'Regan and Sahu [16] introduced the S-iteration method whose algorithmic structure is independent of both Mann and Ishikawa iteration and converge at a faster rate. In last decade, S-iteration method has great attention of researchers as alternative iteration method for solving fixed-point problems and many real-world problems (see, [13, 17–20]). In 2011, Sahu [20] given another variant of S-iteration, which named as normal S-iteration (1.3) which is defined as follows:

$$u_{n+1} = \mathcal{T}[(1 - \alpha_n)u_n + \alpha_n\mathcal{T}u_n] \quad \text{for all } n \in \mathbb{N}, \tag{1.3}$$

where sequence $\{\alpha_n\}$ in $(0, 1)$ is iteration parameter. Both normal S-iteration and S-iteration are two-step processes. They have better performance than Picard (1.1) and Mann iteration method (1.2) for contraction mappings.

Motivated and inspired by work of Yang and Ruan [3], Mann [14] and Sahu [20], we introduce Mann ILC method and normal S-ILC method and study their convergence properties for a kind of discrete time-invariant linear switched systems coupled with an arbitrary switching rule by utilizing super-vector approach, when the system is noise-free. We give a numerical experiment to compare the convergence behaviour proposed ILC methods with P-type ILC method.

The rest of the paper is organized as follows. In Sect. 2, problem formulation and its super-vector representation are described. In Sect. 3, some ILC methods based on Mann and normal S-iteration with their convergence analysis are given. With an error table, a simulation example is given in Sect. 4. Conclusions are drawn in Sect. 5.

2 Preliminaries

In this section, some notations, representations, lemmas and theorems which play a significant role in convergence analysis of the main result have been summarized.

Following notations are used throughout this paper:

- $x_{n+1}(k) \in \mathbb{R}^m$ denote state variable;
- $u_{n+1}(k) \in \mathbb{R}$ denote input signal;
- $y_{n+1}(k + 1) \in \mathbb{R}$ denote output signal;
- The subscript $n \in \mathbb{N}$ denote iteration numbers;
- $P = \{1, 2, 3, \dots, p\}$ with $p \in \mathbb{Z}_+, p < +\infty$;
- $I = \{0, 1, 2, \dots, N_0 - 1\}$, $N_0 \in \mathbb{Z}_+$ indicate operation time period of each trial k ; $k \in I$, denote discrete time instant;
- $\sigma(k)$ denote a random switching rule, which is defined by $\sigma(\cdot) : \mathbb{N} \longrightarrow P$.

Let us consider a single-input, single-output linear discrete time-invariant switched systems for described as follows:

$$\begin{cases} x_{n+1}(k + 1) = A_{\sigma(k)}x_{n+1}(k) + B_{\sigma(k)}u_{n+1}(k), \\ y_{n+1}(k + 1) = C_{\sigma(k)}x_{n+1}(k + 1), k \in I. \end{cases} \tag{2.1}$$

As we see $\sigma(k)$ represents a random switching rule, this means the matrix group $(A_{\sigma(k)}, B_{\sigma(k)}, C_{\sigma(k)})$ is an arbitrary element of the set

$$\{(A_1, B_1, C_1), (A_2, B_2, C_2), \dots, (A_p, B_p, C_p)\}.$$

Objective of system (2.1) is to generate the input sequence $\{u_n(k), k \in I, n \in \mathbb{Z}_+\}$ for which $\{y_n(k + 1), k \in I, n \in \mathbb{Z}_+\}$ converges to reference trajectory $\{y_d(k + 1), k \in I\}$ as closely as possible, mathematically:

$$\lim_{n \rightarrow \infty} |e_{n+1}(k + 1)| = 0 \text{ for all } k \in I,$$

where $e_{n+1}(k + 1) = y_d(k + 1) - y_n(k + 1)$ is the tracking error.

Within present paper, we have considered the following basic assumptions as follows:

Assumption – I (i): All operators begin at identical initial state, i.e., $x_{n+1}(0) = x_0$, for all $n \in \mathbb{Z}_+$ WLOG, assumed that $x_0 = 0$.

(ii) The reference trajectory $y_d(k + 1), k \in I$ is iteration invariant.

Assumption – II : For a given reference trajectory $y_d(k + 1), k \in I$, there exists a reference control input signal $u_d(k), k \in I$ and a reference state vector $x_d(k), k \in I$ such that

$$\begin{cases} x_d(k + 1) = A_{\sigma(k)}x_d(k) + B_{\sigma(k)}u_d(k), \\ y_d(k + 1) = C_{\sigma(k)}x_d(k + 1), k \in I. \end{cases}$$

The P -type ILC method introduced in [3] as follows:

$$u_{n+1}(k) = u_n(k) + \tau_{\sigma(k)}e_n(k + 1), \tag{2.2}$$

where $\tau_{\sigma(k)}$ denote proportion learning gain.

Super-vector representation of ILC [3]. The learning performance of the method (2.2) for the switch system (2.1) by the super-vector approach for given arbitrary time instants $k \in I$, of the $(n + 1)^{th}$ trial given by:

$$Y_{n+1} = H(k)U_{n+1} + D(k), \tag{2.3}$$

where

- $U_{n+1} = [u_{n+1}(0), u_{n+1}(1), \dots, u_{n+1}(k)]^T$
- $Y_{n+1} = [y_{n+1}(1), y_{n+1}(2), \dots, y_{n+1}(k + 1)]^T$
- $H(k) = \begin{bmatrix} C_{\sigma(0)}B_{\sigma(0)} & 0 & 0 & \dots & 0 & 0 \\ C_{\sigma(1)}A_{\sigma(1)}B_{\sigma(0)} & C_{\sigma(1)}B_{\sigma(1)} & 0 & \dots & 0 & 0 \\ C_{\sigma(2)}A_{\sigma(2)}A_{\sigma(1)}B_{\sigma(0)} & C_{\sigma(2)}A_{\sigma(2)}B_{\sigma(1)} & C_{\sigma(2)}B_{\sigma(2)} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{\sigma(k)}\prod_{i=1}^k A_{\sigma(i)}B_{\sigma(0)} & C_{\sigma(k)}\prod_{i=2}^k A_{\sigma(i)}B_{\sigma(1)} & C_{\sigma(k)}\prod_{i=3}^k A_{\sigma(i)}B_{\sigma(2)} & \dots & C_{\sigma(k)}A_{\sigma(k)}B_{\sigma(k-1)} & C_{\sigma(k)}A_{\sigma(k)} \end{bmatrix}$
- $D(k) = [C_{\sigma(0)}A_{\sigma(0)}x_{n+1}(0), C_{\sigma(1)}\prod_{i=0}^1 A_{\sigma(i)}x_{n+1}(0), C_{\sigma(2)}\prod_{i=0}^2 A_{\sigma(i)}x_{n+1}(0), \dots, C_{\sigma(k)}\prod_{i=0}^k A_{\sigma(i)}x_{n+1}(0)]$.

We can deduce that $D(k) = 0$, from Assumption-I. Taking $k = N_0 - 1$, and denoting $H(N_0 - 1) = H$, the system (2.3) becomes

$$Y_{n+1} = HU_{n+1}, \tag{2.4}$$

where $U_{n+1} = [u_{n+1}(0), u_{n+1}(1), \dots, u_{n+1}(N_0 - 1)]^T$ and $Y_{n+1} = [y_{n+1}(1), y_{n+1}(2), \dots, y_{n+1}(N_0)]^T$.

For detail information please check [3].

Accordingly, in the form of super-vector, the control objective of ILC can be similarly described as to generate the super-vector $\{U_n, n \in \mathbb{Z}_+\}$ for which $\{Y_n, n \in \mathbb{Z}_+\}$ converges to a reference trajectory Y_d as closely as possible, i.e.

$$\lim_{n \rightarrow \infty} |e_{n+1}| = 0,$$

where “ $\|\cdot\|$ ” denotes some vector norm and e_{n+1} stand for tracking errors defined as:

$$e_{n+1} = Y_d - Y_{n+1} = [e_{n+1}(1), e_{n+1}(2), \dots, e_{n+1}(N_0)]^T, \tag{2.5}$$

with $Y_d = [y_d(1), y_d(2), \dots, y_d(N_0)]^T$.

Update law (2.2) can be restated in the form of super-vector as follows:

$$U_{n+1} = U_n + \Gamma e_n, \tag{2.6}$$

where

$$\Gamma = \begin{bmatrix} \tau_{\sigma(0)} & 0 & 0 & \cdots & 0 \\ 0 & \tau_{\sigma(1)} & 0 & \cdots & 0 \\ 0 & 0 & \tau_{\sigma(2)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \tau_{\sigma(N-1)} \end{bmatrix} \tag{2.7}$$

is matrix representation of learning gains by updating law (2.2), which is known as learning gain matrix. Now original 2-D ILC problem is transformed into a 1-D linear input–output feedback problem.

Now, by utilizing (2.5) in (2.6), we have:

$$\begin{aligned} U_{n+1} &= U_n + \Gamma e_n \\ &= U_n + \Gamma(Y_d - Y_n) \\ &= U_n + \Gamma Y_d - \Gamma H U_n \\ &= (I - \Gamma H)U_n + \Gamma Y_d, \end{aligned} \tag{2.8}$$

We now define an operator $\mathcal{T} : \mathbb{R}^{N_0} \longrightarrow \mathbb{R}^{N_0}$ by

$$\mathcal{T}(U) = (I - \Gamma H)U + \Gamma Y_d. \tag{2.9}$$

Thus, (2.8) restated as:

$$U_{n+1} = \mathcal{T}U_n, \tag{2.10}$$

which is the useful form of P -type ILC method ([3]).

Throughout the paper, we denote $\|I - H\Gamma\|$ by κ , where “ $\|\cdot\|$ ” denotes the induced matrix norm and \mathcal{T} is defined by (2.9).

Remark 2.1 We see from (2.7) that Γ is a diagonal matrix. Hence, $\Gamma H = H\Gamma$ and hence $\kappa = \|I - H\Gamma\| = \|I - \Gamma H\|$.

The following two theorems are related to the super-vector representation of the system.

Theorem 2.1 ([3], Theorem 1) *Suppose that the ILC method (2.6) is imposed on the switched system (2.4). Then the tracking error converges monotonically if the inequality*

$$\kappa < 1 \tag{2.11}$$

holds.

In consideration of measurement disturbance $\psi_{n+1}(k + 1)$, $k \in I$, the system (2.1) can be restated as

$$\begin{cases} x_{n+1}(k+1) = A_{\sigma(k)}x_{n+1}(k) + B_{\sigma(k)}u_{n+1}(k), \\ y_{n+1}(k+1) = C_{\sigma(k)}x_{n+1}(k+1) + \psi_{n+1}(k+1), \end{cases} k \in I. \quad (2.12)$$

Thus the system (2.4) can restated as

$$Y_{n+1} = HU_{n+1} + \psi_{n+1}, \quad (2.13)$$

where

- $\psi_{n+1} = [\psi_{n+1}(1), \psi_{n+1}(2), \dots, \psi_{n+1}(N_0)]^T$ denote the measurements noise vector with $\|\psi_{n+1}\| < \infty$.

On basis of measurement disturbance, we have two types of tracking errors:

- uncontaminated tracking error e_n given by $e_{n+1} = Y_d - HU_{n+1}$;
- contaminated tracking error \bar{e}_n given by $\bar{e}_{n+1} = Y_d - HU_{n+1} - \psi_{n+1}$.

We can clearly see the relationship between e_{n+1} and \bar{e}_{n+1} , as below:

$$e_{n+1} = \bar{e}_{n+1} + \psi_{n+1}.$$

So, the updating law (2.6) is given by

$$U_{n+1} = U_n + \Gamma\bar{e}_n, \quad (2.14)$$

by the definition of tracking error vector, uncontaminated tracking error can be evolved as

$$e_{n+1} = (I - H\Gamma)e_n + H\Gamma\psi_n. \quad (2.15)$$

Theorem 2.2 ([3], Theorem 2) Consider (2.15) which is derived from the switched system (2.13) by utilizing updating law (2.14). Then $\lim_{n \rightarrow \infty} \sup \|e_{n+1}\|_2$ is bounded if $\|\psi_n\|_2 = \beta < \infty$ for any $n \in \mathbb{Z}_+$ and $\gamma_{\max}(\kappa) = \bar{\gamma} < 1$, where “sup” represents the supremum of a sequence.

3 Some Novel ILC Methods and Their Convergence Analysis

Within this section, we propose two new ILC methods, Mann iteration-based ILC method and normal S-iteration-based ILC method for single-input, single-output linear discrete time-invariant switched systems and analyse their convergence properties.

3.1 Mann Iterative Learning Control Method and Its Convergence Analysis

Based on the Mann iteration method (1.2) and the ILC update method (2.10), we introduce the following ILC method:

$$U_{n+1} = ((1 - \alpha)I + \alpha T)U_n, \tag{3.1}$$

where T is an operator defined by (2.9) and α is a real number with $0 < \alpha < 1$. We say (3.1) is Mann iterative learning control method (Mann-ILC method).

Theorem 3.1 *The Mann-ILC method (3.1) converges if the inequality (2.11) holds.*

Proof Suppose that $\kappa < 1$. Define operator $F : \mathbb{R}^{N_0} \rightarrow \mathbb{R}^{N_0}$ such that

$$F(U) = ((1 - \alpha)I + \alpha T)U. \tag{3.2}$$

For $U_1, U_2 \in \mathbb{R}^{N_0}$, we have

$$\begin{aligned} & \|F(U_1) - F(U_2)\| \\ &= \|[(1 - \alpha)U_1 + \alpha T U_1] - [(1 - \alpha)U_2 + \alpha T U_2]\| \\ &= \|[(1 - \alpha)U_1 + \alpha(I - \Gamma H)U_1 + \alpha \Gamma Y_d] - [(1 - \alpha)U_2 + \alpha(I - \Gamma H)U_2 + \alpha \Gamma Y_d]\| \\ &= \|(1 - \alpha)U_1 + \alpha(I - \Gamma H)U_1 - (1 - \alpha)U_2 - \alpha(I - \Gamma H)U_2\| \\ &= \|[(1 - \alpha)(U_1 - U_2) + \alpha(I - \Gamma H)(U_1 - U_2)]\| \\ &= \|[(1 - \alpha)I + \alpha(I - \Gamma H)](U_1 - U_2)\| \\ &\leq \|I - \alpha \Gamma H\| \|U_1 - U_2\|. \end{aligned}$$

Observe that

$$\begin{aligned} \|I - \alpha \Gamma H\| &= \|(1 - \alpha)I + \alpha(I - \Gamma H)\| \\ &\leq (1 - \alpha) + \alpha \|I - \Gamma H\| \\ &\leq (1 - \alpha) + \alpha \kappa. \end{aligned} \tag{3.3}$$

Since $\kappa < 1$, the operator F defined by (3.2) is contraction. Therefore, input super-vector sequence $\{U_n, n \in \mathbb{Z}_+\}$ defined by Mann-ILC method (3.1) is convergent.

We now study asymptotic convergence of tracking error of Mann-ILC method (3.1).

Theorem 3.2 *Assume that the Mann-ILC method (3.1) is imposed on switched system (2.4). Then the tracking error converges monotonically if the inequality (2.11) holds.*

Proof Suppose that $\kappa < 1$. By definition of tracking error (2.5), we have

$$e_{n+1} = Y_d - Y_{n+1}.$$

From (3.1) and (2.4), we have

$$\begin{aligned}
 e_{n+1} &= Y_d - HU_{n+1} \\
 &= Y_d - H[(1 - \alpha)I + \alpha\mathcal{T}]U_n \\
 &= Y_d - H[(1 - \alpha)U_n + \alpha(I - \Gamma H)U_n + \alpha\Gamma Y_d] \\
 &= Y_d - H[(I - \alpha\Gamma H)U_n + \alpha\Gamma Y_d] \\
 &= (I - \alpha H\Gamma)Y_d - H(I - \alpha\Gamma H)U_n \\
 &= (I - \alpha H\Gamma)Y_d - (I - \alpha H\Gamma)HU_n \\
 &= (I - \alpha H\Gamma)(Y_d - Y_n) \\
 &= (I - \alpha H\Gamma)e_n.
 \end{aligned}
 \tag{3.4}$$

Taking the norm of both sides (3.4) and utilizing norm inequality, we yield

$$\|e_{n+1}\| \leq \|I - \alpha H\Gamma\| \|e_n\|.$$

Since $\kappa < 1$, it follows, from (3.3), that $\|I - \alpha H\Gamma\| < 1$. Thus, we get

$$\|e_{n+1}\| < \|e_n\|.$$

This completes the proof.

Remark 3.1 For monotone convergence of tracking error condition (2.11) is sufficient. Especially, in the case of $\|\cdot\|_2$.

3.2 Normal S-Iterative Learning Control Method and Its Convergence Analysis

Based on normal S-iteration method (1.3) and the update method (2.10), we define S-operator generated by α and \mathcal{T} .

Definition 3.1 Let α be a real number with $0 < \alpha < 1$ and \mathcal{T} an operator defined by (2.9). Then an operator $G_{\alpha,\mathcal{T}} : \mathbb{R}^{N_0} \rightarrow \mathbb{R}^{N_0}$ is said to be an S-operator generated by α and \mathcal{T} if

$$G_{\alpha,\mathcal{T}} = \mathcal{T}[(1 - \alpha)I + \alpha\mathcal{T}]. \tag{3.5}$$

We give basic properties of S-operator $G_{\alpha,\mathcal{T}}$.

Proposition 3.1 The S-operator $G_{\alpha,\mathcal{T}}$ defined by (3.5) satisfies the following:

1. $G_{\alpha,\mathcal{T}}(U) = (I - \Gamma H)(I - \alpha\Gamma H)U + (I + \alpha(I - \Gamma H))\Gamma Y_d$ for all $U \in \mathbb{R}^{N_0}$.
2. The operator $G_{\alpha,\mathcal{T}}$ is contraction if the inequality (2.11) holds.

Proof 1. For $U \in \mathbb{R}^{N_0}$, we have

$$\begin{aligned}
 G_{\alpha, \mathcal{T}}(U) &= \mathcal{T}[(1 - \alpha)U + \alpha \mathcal{T}U] \\
 &= \mathcal{T}[(1 - \alpha)U + \alpha(I - \Gamma H)U + \alpha \Gamma Y_d] \\
 &= \mathcal{T}[(I - \alpha \Gamma H)U + \alpha \Gamma Y_d] \\
 &= (I - \Gamma H)[(I - \alpha \Gamma H)U + \alpha \Gamma Y_d] + \Gamma Y_d \\
 &= (I - \Gamma H)(I - \alpha \Gamma H)U + (I + \alpha(I - \Gamma H))\Gamma Y_d. \tag{3.6}
 \end{aligned}$$

2. Suppose that inequality (2.11) holds. Let $U_1, U_2 \in \mathbb{R}^{N_0}$. Then, for $V_1 = (1 - \alpha)U_1 + \alpha \mathcal{T}U_1$, we have

$$\begin{aligned}
 G_{\alpha, \mathcal{T}}(U_1) &= \mathcal{T}[V_1] \\
 &= (I - \Gamma H)V_1 + \Gamma Y_d \\
 &= (I - \Gamma H)[(1 - \alpha)U_1 + \alpha \mathcal{T}(U_1)] + \Gamma Y_d \\
 &= (1 - \alpha)(I - \Gamma H)U_1 + \alpha(I - \Gamma H)\mathcal{T}(U_1) + \Gamma Y_d
 \end{aligned}$$

and hence

$$\begin{aligned}
 &\|G_{\alpha, \mathcal{T}}(U_1) - G_{\alpha, \mathcal{T}}(U_2)\| \\
 &= \|(1 - \alpha)(I - \Gamma H)U_1 + \alpha(I - \Gamma H)\mathcal{T}U_1 + \Gamma Y_d - [(1 - \alpha)(I - \Gamma H)U_2 + \alpha(I - \Gamma H)\mathcal{T}U_2 + \Gamma Y_d]\| \\
 &= \|(1 - \alpha)(I - \Gamma H)(U_1 - U_2) + \alpha(I - \Gamma H)[\mathcal{T}U_1 - \mathcal{T}U_2]\| \\
 &= \|(1 - \alpha)(I - \Gamma H)(U_1 - U_2) + \alpha(I - \Gamma H)(I - \Gamma H)(U_1 - U_2)\| \\
 &= \|(1 - \alpha)(I - \Gamma H) + \alpha(I - \Gamma H)(I - \Gamma H)\|(U_1 - U_2)\| \\
 &= \|(I - \Gamma H)(I - \alpha \Gamma H)\|(U_1 - U_2)\| \\
 &\leq \|(I - \Gamma H)(I - \alpha \Gamma H)\| \|U_1 - U_2\|.
 \end{aligned}$$

Observe that

$$\begin{aligned}
 \|(I - \Gamma H)(I - \alpha \Gamma H)\| &= \|(1 - \alpha)(I - \Gamma H) + \alpha(I - \Gamma H)(I - \Gamma H)\| \\
 &\leq (1 - \alpha)\|I - \Gamma H\| + \alpha\|(I - \Gamma H)(I - \Gamma H)\| \\
 &\leq (1 - \alpha)\kappa + \alpha\kappa^2.
 \end{aligned}$$

Since $\kappa < 1$, it follows that $\|(I - \Gamma H)(I - \alpha \Gamma H)\| < 1$. Therefore, the operator $G_{\alpha, \mathcal{T}}$ defined by (3.2) is contraction.

In view of normal S-iteration method (1.3), we introduce the following ILC method:

$$U_{n+1} = G_{\alpha, \mathcal{T}}U_n, \tag{3.7}$$

where T is an operator defined by (2.9) and α is a real number with $0 < \alpha < 1$. We say (3.7) is normal S-iterative learning control method (normal S-ILC method).

By utilizing (3.6), the normal S-ILC method (3.7) can be restated as

$$U_{n+1} = (I - \Gamma H)(I - \alpha \Gamma H)U_n + (I + \alpha(I - \Gamma H))\Gamma Y_d. \quad (3.8)$$

Theorem 3.3 *The normal S-ILC method (3.7) converges if inequality (2.11) holds.*

Proof Suppose that inequality (2.11) holds. From proposition 3.1, we see that the operator $G_{\alpha, \mathcal{T}}$ defined by (3.5) is contraction. Therefore, input super-vector sequence $\{U_n, n \in \mathbb{Z}_+\}$ defined by normal S-ILC method (3.7) is convergent.

We now study asymptotic convergence of tracking error of normal S-ILC algorithm (3.7).

Theorem 3.4 *Assume that the normal S-ILC method (3.7) is imposed on the switched system (2.4). Then the tracking error converges monotonically if the inequality (2.11) holds.*

Proof By definition of tracking error (2.5), we have

$$e_{n+1} = Y_d - Y_{n+1}.$$

From (2.4) and (3.8), we have

$$\begin{aligned} e_{n+1} &= Y_d - HU_{n+1} \\ &= Y_d - H((I - \Gamma H)(I - \alpha \Gamma H)U_n + (I + \alpha(I - \Gamma H))\Gamma Y_d) \\ &= (I - H\Gamma - \alpha H\Gamma + \alpha H\Gamma H\Gamma)Y_d - H((I - \Gamma H)(I - \alpha \Gamma H)U_n \\ &= (I - H\Gamma)(I - \alpha H\Gamma)Y_d - (I - \Gamma H)(I - \alpha \Gamma H)HU_n \\ &= (I - H\Gamma)(I - \alpha H\Gamma)Y_d - (I - H\Gamma)(I - \alpha H\Gamma)HU_n \\ &= (I - H\Gamma)(I - \alpha H\Gamma)(Y_d - Y_n) \\ &= (I - H\Gamma)(I - \alpha H\Gamma)e_n. \end{aligned} \quad (3.9)$$

By taking the norm of both sides (3.9) and utilizing norm inequality, we yield

$$\|e_{n+1}\| \leq \|(I - H\Gamma)(I - \alpha H\Gamma)\| \|e_n\|.$$

If the assumption (2.11) holds, we get

$$\|e_{n+1}\| < \|e_n\|.$$

This completes the proof.

4 Simulation Example

In order to express the efficiency of the proposed ILC methods, we presented a simulation example. We consider a linear discrete-time switched system coupled with three subsystems given in [3] as follows :

$$\begin{cases} x_{n+1}(k+1) = A_{\sigma(k)}x_{n+1}(k) + B_{\sigma(k)}u_{n+1}(k), \\ y_{n+1}(k+1) = C_{\sigma(k)}x_{n+1}(k+1), \quad k \in I. \end{cases}$$

For given arbitrary v_k belongs to $\{1, 2, \dots, 60\}$, arbitrary switching sequence is defined as below:

$$\sigma(k) = \begin{cases} 1, & \text{if } v_k \bmod 3 = 1, \\ 2, & \text{if } v_k \bmod 3 = 2, \\ 3, & \text{if } v_k \bmod 3 = 0. \end{cases}$$

For simplicity we take $I = \{0, 1, 2, \dots, 59\}$. Initial condition:

- $u_1(k) = 0$, for $k \in I$;
- $x_{n+1}(0) = 0$ for all $n \in \mathbb{N}$.

The system dynamics are given by

$$\begin{aligned} A_1 &= \begin{bmatrix} 0 & 1 \\ 0.125 & -0.2 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C_1 = [0.1 \ 1]; \\ A_2 &= \begin{bmatrix} -0.25 & 1 \\ 0 & -0.3 \end{bmatrix}, B_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C_2 = [-0.2 \ 1]; \\ A_3 &= \begin{bmatrix} 1 & 0 \\ 0.2 & -0.1 \end{bmatrix}, B_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C_3 = [0.25 \ 1]. \end{aligned}$$

The reference trajectory is provided as

$$y_d(k+1) = \sin\left(\frac{8}{25}k\right), k \in I. \quad (4.1)$$

In our experiment, we take $\alpha = 0.9$ for Mann-ILC and normal S-ILC method. One of the possible switching sequences $\sigma(k)$ is given as

$$\sigma(k) = \begin{cases} 1, & \text{if } k \bmod 3 = 1, \\ 2, & \text{if } k \bmod 3 = 2, \\ 3, & \text{if } k \bmod 3 = 0. \end{cases}$$

As shown in Fig. 2. We consider notations as follows:

1. Y_p and e_p denote output and output error for P-type (2.10) method respectively;
2. Y_m and e_m denote output and output error for Man-ILC method (3.1), respectively;

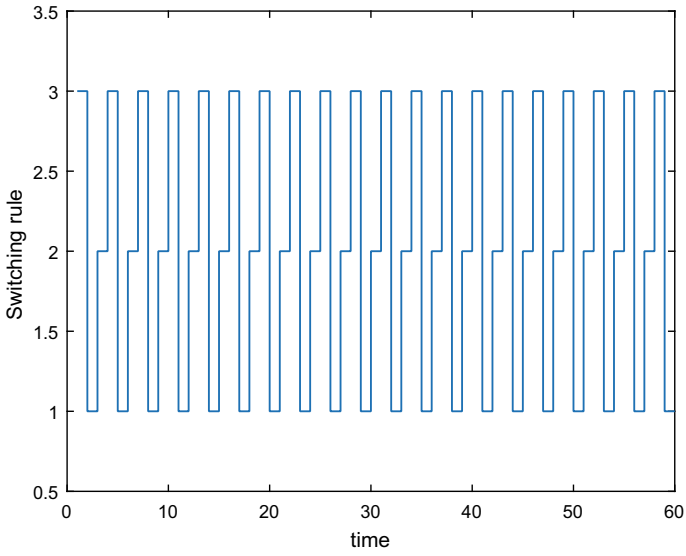


Fig. 2 Switching sequence $\sigma(k)$

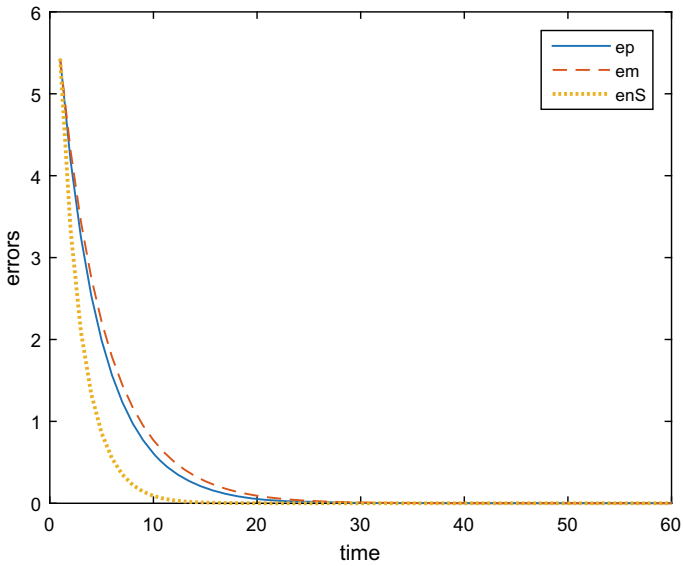


Fig. 3 Coordinatewise graph of error for different iteration methods

- 3. Y_{nS} and e_{nS} denote output and output error for normal S-ILC method (3.8), respectively.

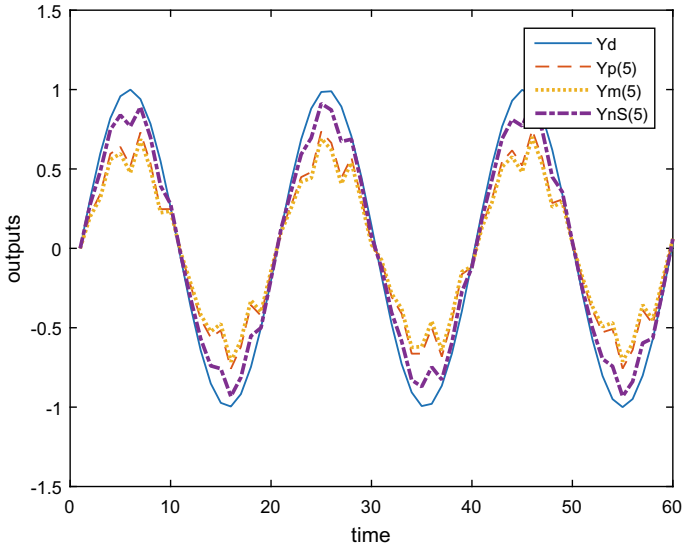


Fig. 4 System outputs of P-type ILC method, Mann-ILC method and normal S-ILC method at 5th trial

The system outputs of P-type method (2.10), Mann-ILC method (3.1) and normal S-ILC method (3.8) at 5th trial, the 15th trial and the 25th trial are shown in Figs. 4, 5 and 6, respectively.

For proportion learning gain $\tau = 0.25$, error ($e_n, n \in \mathbb{N}$) corresponding to the P-type ILC method (2.10), Mann-ILC method (3.1) and normal S-ILC method (3.8) are given in Table 1 and Fig. 3, in which we can see that error corresponding to P-type ILC method, Mann-ILC method and normal S-ILC method converges at 46th, 52nd and 25th iteration, respectively.

5 Conclusions

In this paper, Mann (1.2) and normal S-iteration (1.3) are applied to the super-vector representation of a kind of discrete switched systems coupled with an arbitrary switching rule. Some methods are presented for such a switched system with corresponding convergence condition. We show that under suitable condition, on the whole interval the method can ensure the output tracking error converges to zero. At last, the stimulation result demonstrates the efficiency of the presented algorithms and it manifests that system converges faster for normal S-ILC method (3.8) than P-type ILC method (2.10) and Mann ILC method (3.1) but converges slower for Mann ILC algorithm than P-type ILC method.

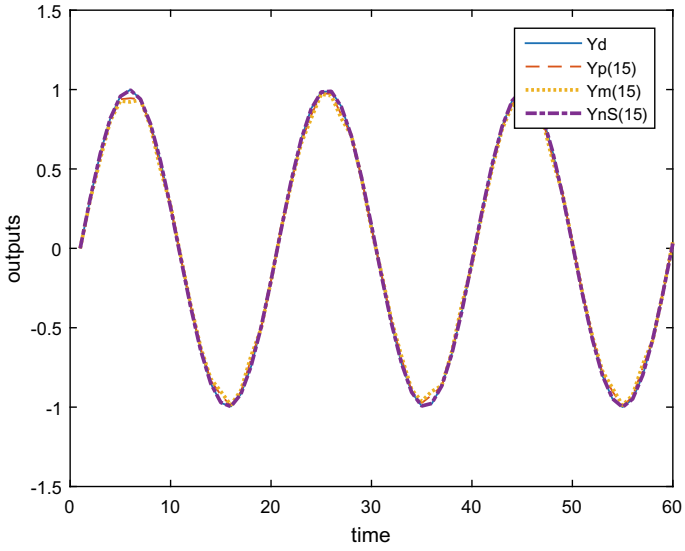


Fig. 5 System outputs of P-type ILC method, Mann-ILC method and normal S-ILC method at 15th trial

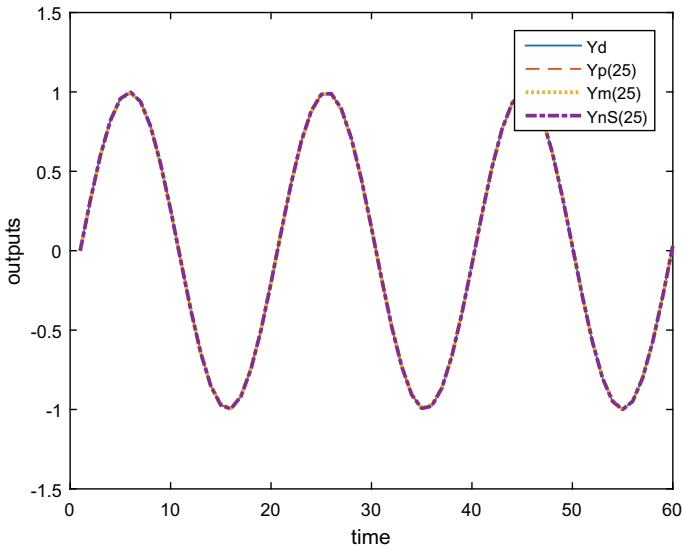


Fig. 6 System outputs of P-type ILC method, Mann-ILC method and normal S-ILC method at 25th trial

Table 1 Error table

No. of iteration	P-type	Mann	Normal-S	No. of iteration	P-type	Mann	Normal-S
1	5.4272	5.4272	5.4272	29	0.0055	0.0124	0.0000
2	4.1878	4.3103	3.3436	30	0.0043	0.0099	0.0000
3	3.2510	3.4392	2.0993	31	0.0033	0.0079	0.0000
4	2.5373	2.7557	1.3353	32	0.0025	0.0063	0.0000
5	1.9891	2.2162	0.8552	33	0.0020	0.0050	0.0000
6	1.5648	1.7877	0.5486	34	0.0015	0.0040	0.0000
7	1.2341	1.4454	0.3513	35	0.0012	0.0031	0.0000
8	0.9748	1.1706	0.2241	36	0.0009	0.0025	0.0000
9	0.7706	0.9492	0.1422	37	0.0007	0.0020	0.0000
10	0.6093	0.7701	0.0897	38	0.0005	0.0016	0.0000
11	0.4816	0.6249	0.0563	39	0.0004	0.0012	0.0000
12	0.3803	0.5070	0.0351	40	0.0003	0.0010	0.0000
13	0.3000	0.4111	0.0218	41	0.0002	0.0008	0.0000
14	0.2364	0.3331	0.0135	42	0.0002	0.0006	0.0000
15	0.1859	0.2697	0.0083	43	0.0001	0.0005	0.0000
16	0.1460	0.2181	0.0051	44	0.0001	0.0004	0.0000
17	0.1145	0.1762	0.0031	45	0.0001	0.0003	0.0000
18	0.0896	0.1422	0.0019	46	0.0001	0.0002	0.0000
19	0.0701	0.1146	0.0012	47	0.0000	0.0002	0.0000
20	0.0547	0.0922	0.0007	48	0.0000	0.0002	0.0000
21	0.0426	0.0741	0.0004	49	0.0000	0.0001	0.0000
22	0.0332	0.0595	0.0003	50	0.0000	0.0001	0.0000
23	0.0258	0.0477	0.0002	51	0.0000	0.0001	0.0000
24	0.0200	0.0382	0.0001	52	0.0000	0.0001	0.0000
25	0.0155	0.0306	0.0001	53	0.0000	0.0000	0.0000
26	0.0120	0.0245	0.0000	54	0.0000	0.0000	0.0000
27	0.0093	0.0195	0.0000	55	0.0000	0.0000	0.0000
28	0.0072	0.0156	0.0000				

References

1. B. Clymer, S.A. Collins, Optical computer switching network. *Optical Eng.* **24**(1), 240174 (1985)
2. Y.B. Kim, S.S. Lee, C.H. Oh, Y.S. Kim, C. Han, C.H. Yim, An architecture of scalable ATM switching system and its call processing capacity estimation. *ETRI J.* **18**(3), 107–125 (1996)
3. X. Yang, X. Ruan, Analysis of iterative learning control for a class of linear discrete-time switched systems. *Abst. Appl. Anal.* **1–8**, (2015)
4. X.H. Bu, F.S. Yu, Z.S. Hou, F.Z. Wang, Iterative learning control for a class of linear discrete-time switched systems. *Acta Autom. Sinica* **39**(9), 1564–1569 (2013)

5. T. Weller, B. Hajek, Scheduling nonuniform traffic in a packet switching system with small propagation delay. *IEEE/ACM Trans. Networking* **5**(6), 558–598 (1998)
6. H. Lin, P.J. Antsaklis, Stability and stabilizability of switched linear systems: a survey of recent results. *IEEE Trans. Autom. Control* **54**(2), 308–322 (2009)
7. Z. Ji, L. Wang, X. Guo, Design of switching sequences for controllability realization of switched linear systems. *Automatica* **43**(4), 662–668 (2007)
8. J.P. Hespanha, D. Liberzon, D. Angeli, E.D. Sontag, Nonlinear norm-observability notions and stability of switched systems. *IEEE Trans. Autom. Control* **50**(2), 154–168 (2005)
9. H.S. Ahn, Y.Q. Chen, K.L. Moore, Iterative learning control: brief survey and categorization. *IEEE Trans. Syst. Man Cybern.* **37**, 1099–1121 (2007)
10. S. Arimoto, S. Kawamura, F. Miyazaki, Bettering operation of robots by learning. *J. Robot. Syst.* **1**(2), 123–140 (1984)
11. K.L. Moore, Y.Q. Chen, H.S. Ahn, Iterative learning control: a tutorial and big picture view, in *Proceedings of 45th IEEE Conference of Decision and Control* (2006), pp. 2352–2357
12. J.J.A. Eksteen, P.S. Heyns, An alternative update formula for non-linear model-based iterative learning control. *Inverse Probl. Sci. Eng.* **24**(5), 860–888 (2016)
13. R.P. Agarwal, D. O'Regan, D.R. Sahu, *Fixed Point Theory for Lipschitzian-Type Mappings with Applications, Series: Topological Fixed Point Theory and Its Applications*, 6 (Springer, New York, 2009)
14. W.R. Mann, Mean value methods in iteration. *Proc. Amer. Math. Soc.* **4**, 506–610 (1953)
15. S. Ishikawa, Fixed points by a new iteration method. *Proc. Amer. Math. Soc.* **44**, 147–150 (1974)
16. R.P. Agarwal, D. O'Regan, D.R. Sahu, Iterative construction of fixed points of nearly asymptotically nonexpansive mappings. *J. Nonlinear Convex Anal.* **8**(1), 61–79 (2007)
17. D.R. Sahu, Applications of accelerated computational methods for quasi-nonexpansive operators to optimization problems. *Soft Comput.* **24**, 17887–17911 (2020). <https://doi.org/10.1007/s00500-020-05038-9>
18. A. Dixit, D.R. Sahu, A.K. Singh, T. Som, Application of a new accelerated algorithm to regression problems. *Soft Comput.* **24**(2), 1539–1552 (2020)
19. F. Gürsoy, J.J.A. Eksteen, A.R. Khan, V. Karakaya, An iterative method and its application to stable inversion. *Soft Comput.* **23**(16), 7393–7406 (2019)
20. D.R. Sahu, Applications of the S-iteration process to constrained minimization problems and split feasibility problems. *Fixed Point Theory and Appl.* **12**, 187–204 (2011)

A Scalable VLSI Architecture for Illumination-Invariant Heterogeneous Face Recognition



Swagata Bhattacharya, Somsubhra Talapatra, Debotosh Bhattacharjee, and Amlan Chakrabarti

Abstract This paper proposes a scalable VLSI architecture of the local gravity face angle (LGFA) feature extraction algorithm. The LGFA is a novel feature extraction algorithm used for generating LG-face for illumination-invariant and heterogeneous face recognition. The proposed approach uses fixed-point arithmetic representation and min-max normalization process is modified at algorithmic and architecture stage for faster average computation (convergence) of angles. The architecture presented in this paper is pipelined and parallel and have no preprocessing stage for zero padding. The design is validated in Xilinx Virtex-5 FPGA (XC5VLX110T). The proposed architecture is scalable for any image up to 128×128 pixels size and the limitation is only due to the size of the internal counter and memory resources (Block RAM) available in FPGA platform. The simulation of the design in Xilinx ISE 14.7 shows that the emulated architecture can compute the LGFA of 128×128 image with variation of 0.27° at 68.312 MHz consuming only 42mW of switching power. This paper is a first attempt in regard to FPGA implementation of LGFA feature extraction.

Keywords Local gravity face angle · Heterogeneous face recognition · Pipelined and paralleled architecture · FPGA implementation

S. Bhattacharya (✉)

Guru Nanak Institute of Technology, Kolkata, West Bengal 700114, India

S. Talapatra

Aliah University, New Town, Kolkata, West Bengal 700160, India

D. Bhattacharjee

Jadavpur University, Kolkata, West Bengal 700032, India

A. Chakrabarti

A.K. Choudhury School of Information Technology, Faculty of Engineering and Technology, University of Calcutta, Kolkata, West Bengal 700009, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

459

S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,

Lecture Notes in Electrical Engineering 756,

https://doi.org/10.1007/978-981-16-0749-3_34

1 Introduction

Face recognition is a very powerful application of image analysis and gained huge application in the field of computer vision. The reliable face recognition still faces many challenges though many reliable systems have been developed [1]. Illumination variation is such a major challenge in face recognition. A wide range of approaches are used to cope with illumination variation problem. These methods are categorized in three broad ways (1) preprocessing [2, 3], (2) modelling [4, 5] and (3) invariant feature representation [6, 7]. A few researchers [8, 9] used images of different modalities to minimize the impact of illumination. To address such images of different modalities, a field of face recognition called heterogeneous face recognition (HFR) is emerging [10].

Roy and Bhattacharjee [11] proposed an illumination-invariant feature extraction method based on the use of direction of the local gravitational force, expressed in the angular form of local gravitational force angle (LGFA). The direction feature is called local gravity face angle and the method is called local gravity face (*LG-face*). This method is equally able to cope with illumination variations and heterogeneous face recognition. It is more adaptive to local variation in illumination and yields a better result than other existing methods (KP-RS [12], LCKS-CSR [13] and the THFM [10]). Moreover, it performs consistently on the LFW database and in the presence of noise. In [14], this LGFA method is used to come up with a novel image descriptor called a pattern of local gravitational forces (PLGF). The hardware implementations of face recognition algorithms are very important for a wide variety of real-time applications like surveillance, biometric verifications (e.g. user authentications in smart devices) and security (e.g. door access control), etc. Face recognition feature in mobile devices enables various customer services. Though many research works have been carried out [15–18] in recent times on developing VLSI architectures of face recognition methods for FPGA/ASIC implementations, the hardware implementations of the illumination-invariant heterogeneous face recognition system are still an area that needs much more exploration. In recent years, some works present accelerator designs for face recognition [19, 20] but since they are neural network based, they are computation intensive. As can be seen from above discussions LGFA method shows prospect in its use for real-time face recognition. The system can be attached with CCTV cameras installed in buildings, streets and in any surveillance zone. Also, software realization of face recognition algorithms in mobile processors consumes a power $> 1W$ as discussed in [19]. This motivates to design a hardware implementation of this algorithm, primarily designing a scalable VLSI architecture for LGFA feature extraction for FPGA implementation. The FPGA platform is chosen due to portability, and it provides reconfigurability and shorter turn around time of the chip with low cost. The purpose of this paper is to present a thorough description of VLSI architecture for LGFA feature extraction and to demonstrate its functionality, implementation in FPGA and performance. The major contributions of this work are summarized below:

Claim

1. This is a first attempt in regard to FPGA-based parallel hardware implementation of LGFA extraction.
2. Any grayscale image of size up to 128×128 can be processed without changing the hardware. The image size is only limited by the resources available in an implementation platform.
3. No preprocessing of an image is required for pixel extensions by zero padding in hardware.
4. Faster average computation (convergence) of angles of local gravity force is achieved using a modified vector mode CORDIC. The proposed architecture was implemented on a Xilinx Virtex-5 FPGA (XC5VLX110T) using ISE 14.5. This emulated architecture was able to compute LGFA of a 128×128 grey-scale image with an error of 0.27° , compared to the MATLAB results, at 68.312 MHz consuming only 42mW of switching power.
5. In this design, fixed-point arithmetic is used which is faster and consumes lesser hardware resources than floating-point arithmetic. The number of bits for fixed-point numbers is chosen so that we do not lose out accuracy.
6. The min-max normalization in LGFA computation [11] involves the division of each LGFA minus the minimum LGFA angle in the LGFA matrix, by the absolute difference between maximum and minimum values in the matrix. To simplify the complexity of division, we modify the algorithm of the min-max normalization and this modification is incorporated into the architecture as well.

In Sect. 2 we give preliminary introduction to LGFA and LG-face. In Sect. 3, we describe in detail (a) the logic parallelism existing in the algorithm of LGFA feature extraction method, (b) the need to exploit the parallelism in FPGA and (c) proposed architecture and design considerations. Experimental results and comparisons are presented in Sect. 4 and finally, the paper is concluded in Sect. 5.

2 Preliminaries

Roy and Bhattacharjee [11] have used the gravitational force angle to analyse large and challenging face databases, primarily for illumination-invariant face recognition and HFR.

A: LG-face Mask: The LG -face Mask [11] is as shown in Fig. 1.

C: LGFA feature extraction for LG-face LGFA is extracted from an image to generate LG-face which is used for face recognition. Here is the algorithm for extracting LGFA feature from an image:

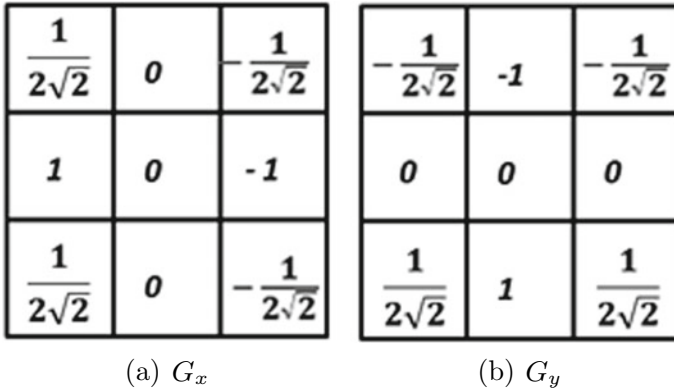


Fig. 1 LGFA horizontal **a** and vertical **b** convolution masks

Algorithm 1 Algorithm to Extract Local Gravitational Force Angle(LGFA) Feature

Input: Sample Image.

Output: LG-face Image.

- 1: Apply Mask G_x, G_y to the sample image.
 - 2: Result of masking with G_x is stored in X
 - 3: Result of masking with G_y is stored in Y
 - 4: Result combined to find the LGFA: $LGFA = \arctan(Y/X)$
 - 5: Perform Min-Max normalization operation on LGFA to get local gravity face image
-

3 Proposed VLSI Architecture for Local Gravity Face Angle (LGFA) Extraction

The block diagram of overall architecture for LGFA computation and verification is shown in Figure 2. It consists of four major blocks: (a) internal memory blocks: (i) M1 (ii) M2 (b) controller (c) datapath(LGFA)Generation (d) datapath for min-max normalization. To exploit the parallelism in the architecture, we have developed it for FPGA implementation.

3.1 Internal Memory

The memory requirement for this architecture is met with Block RAM (BRAM) available on the FPGA chip. The architecture has two memories M1 and M2. M1 is used to store input grey image with maximum dimension 128×128 . The memory size of M1 is 4096×32 bit. In a 32-bit word, since each pixel is represented by 8 bit, four pixels can be stored. The operation of the memory is controlled by a WR/RD

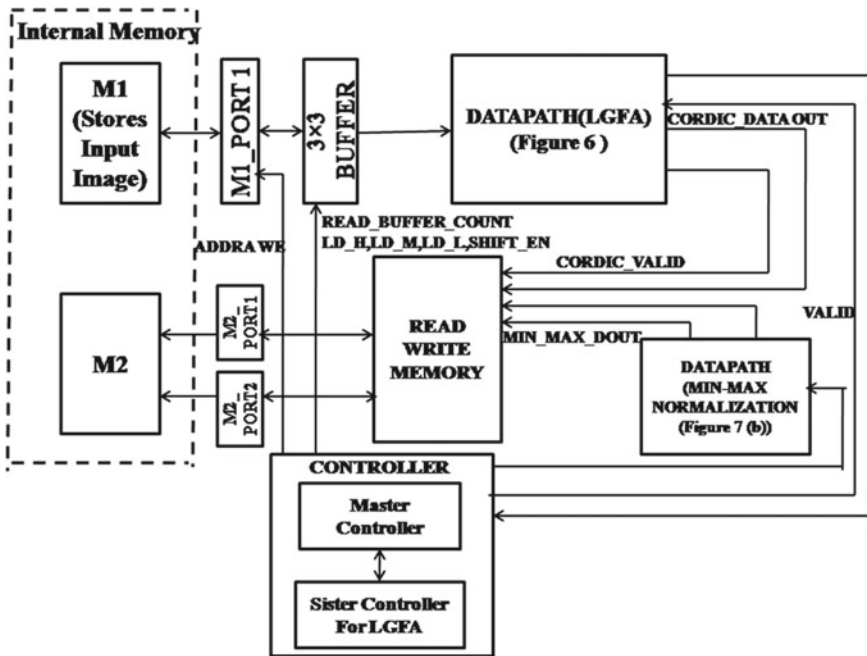


Fig. 2 Block diagram of overall architecture for LGFA computation

signal. It is a single port BRAM (M1_PORT1 as shown in Fig. 2). The memory M2 is configured as $32\text{ K} \times 16$ bit memory. The upper 4096 words are loaded with a golden value of LGFA image computed by MATLAB for the test image. The computed LGFA image computed by hardware is stored in the lower 4096 words of memory M2. This memory arrangement enables to execute the validation phase that will be described under the controller subsection. It is a true dual-port memory (M2 PORT1 and M2 PORT2). The M2 PORT1 supports write accesses from both LGFA datapath as well as min-max datapath. The arbitration between write accesses by two datapath is managed by read and write memory interface.

3.2 Controller

The controller for this architecture consists of: (a) master controller for overall control (b) sister controller for controlling LGFA computation and min-max normalization. The state machine for master controller is shown in Fig. 3. The state machine consists of four states namely: (i) CFG(or configuration state), (ii) LGFA state, (iii) MINMAX and (iv) VALIDATE. CFG: This is a default state of the state machine. LGFA: As start 1/lgfastart signal goes high, LGFA computation state gets active.

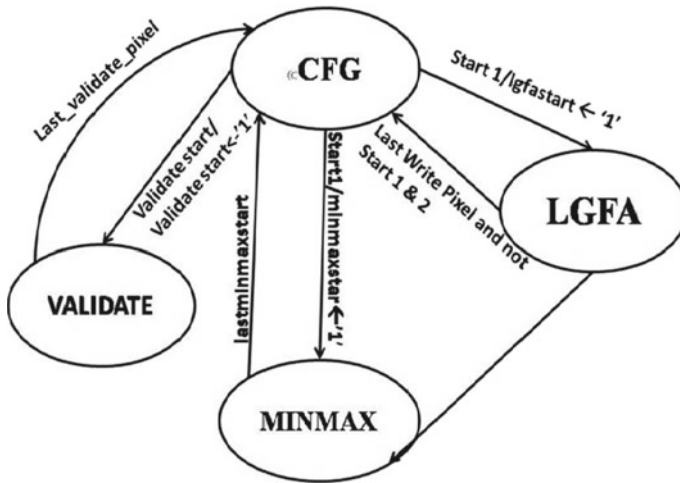


Fig. 3 Master state machine for controller

This is the first stage of LGFA computation. In this state, LGFA image is computed by the hardware (LGFA datapath). When the LGFA state returns the last pixel, Start2 or minmax signal goes high. Then the second phase of computation starts and the controller goes to MINMAX state. MINMAX: In this state, the controller executes the min-max normalization by dividing the angle pixel by 360° using the min-max datapath. The minimum angle pixel for the input image is calculated in the LGFA datapath itself in first phase. For the faster computation of min-max normalization, simultaneous read-write access is made to memory M2. VALIDATION: In validation phase, the computed LGFA angle pixels are validated by executing a verification scheme when 'lastminmax' pixel signal is received in CFG state. The *validate_start* signal goes high and verification starts. As described in the INTERNAL MEMORY section that 4096 words M2 consists of golden LGFA computed by MATLAB and in the lower part of MEMORY, the hardware computed LGFA image is stored. In VALIDATION phase, the controller reads angle pixels from both halves of M2 using two ports and compare numerical equality with some error tolerance and keeps a count of the number of matching pixels. If this matching pixels count is the same as the input image, a signal is flagged as a success. This signal is mapped to FPGA board.

3.3 LGFA Microarchitecture for Datapath

The computations required for all of the LGFA pixels are independent of one another, since they only require values of the original image pixels. Thus, we could perform computations for as many LGFA pixels in parallel as required. For computation

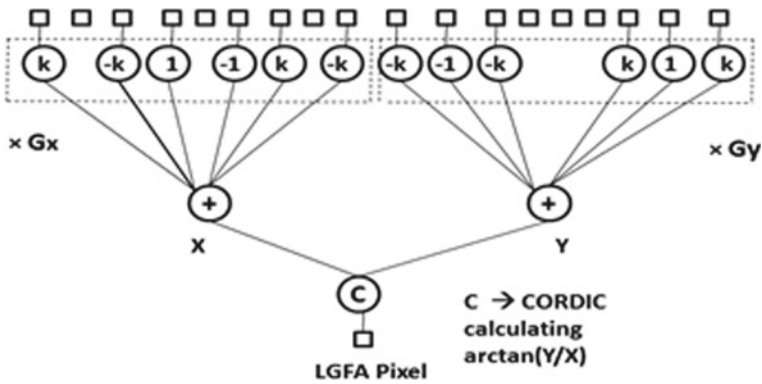


Fig. 4 Data dependency graph for LGFA feature generation

of each LGFA pixel, the data dependency graph is shown in Fig. 4. This diagram shows the data required for each operation, starting with the pixels from the original image at the top, with intermediate results feeding through to dependent operations, yielding the LGFA pixel at the bottom. We have excluded partial products in which the coefficient is 0, since they do not contribute to the result. The diagram shows that we can compute all of the partial products in parallel, since each partial product depends only on an original pixel value and a constant coefficient. As shown in the diagram the multiplier coefficients are $k, -k, 1, -1$ where $k = \frac{1}{2\sqrt{2}}$. We can then sum the two groups of six partial products in parallel, and then calculate the arctan (Y/X) using CORDIC [21, 22] to produce the LGFA pixel value. The parallelism of the graph is exploited in the architecture by the design of LGFA microarchitecture for datapath (Fig. 4).

The algorithm for local gravitational force angle (LGFA) feature extraction method is implemented using the microarchitecture for LGFA as shown in Fig. 5. To implement masking as described in step 2 of Algorithm 1 of Sect. 2, (also shown in data dependency graph of Fig. 4), it is essential to multiply corresponding elements of the mask with the input image. To achieve this process in hardware at first, the LGFA datapath reads four pixels from each of the preceding, current and next rows in memory into the three 32-bit registers at the top right of Fig. 6. Each register consists of four 8-bit pixel registers. To apply G_x and G_y on input 3×3 window, we essentially use a pipeline with 3×3 Buffer (i.e. 8-bit registers each for storing single 8-bit pixel) where pixels are read from the original image entering into the registers at the top right. The pixel then flow through the 3×3 multiplier array. The adders (ADDER_X and ADDER_Y) to the Y and X registers (Y and X are numerator and finally to the CORDIC block that implements $\arctan(Y/X)$ which is LGFA pixel and the result is sent to memory. Coordinate rotation digital computer (CORDIC) is a well-known class of hardware efficient multiplierless algorithms for realization of trigonometric functions [23]. The trigonometric functions are defined over vector rotation and can be efficiently implemented iteratively using shifts and additions with

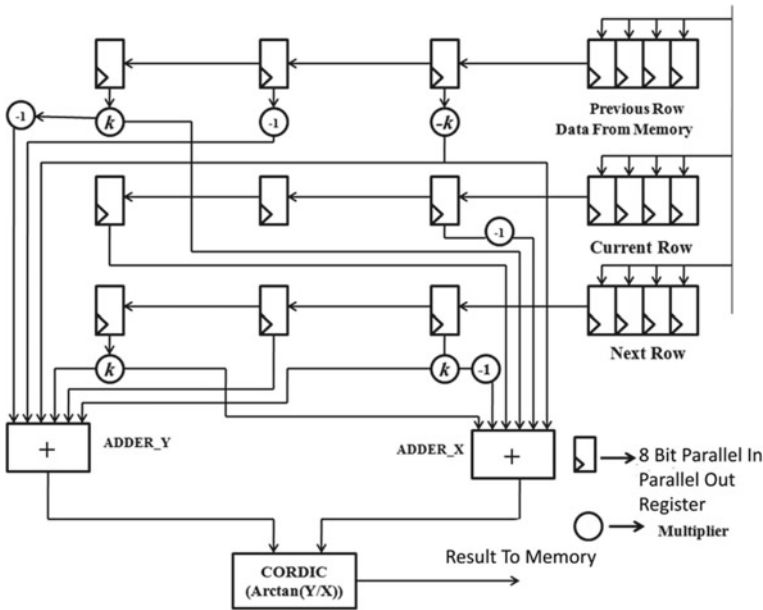


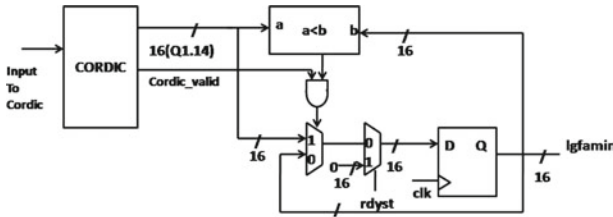
Fig. 5 Microarchitecture of LGFA datapath

the help of a look-up-table of special angles α_i . These angles are selected such that for an i th angle, $\arctan(\alpha_i = 2^{-i})$. Some excellent reviews on CORDIC algorithms and its applications can be found in [21–23]. In the LGFA datapath, a vectoring mode CORDIC algorithm [24] is used to implement the function $\arctan(Y/X)$ as required to implement LGFA feature extraction.

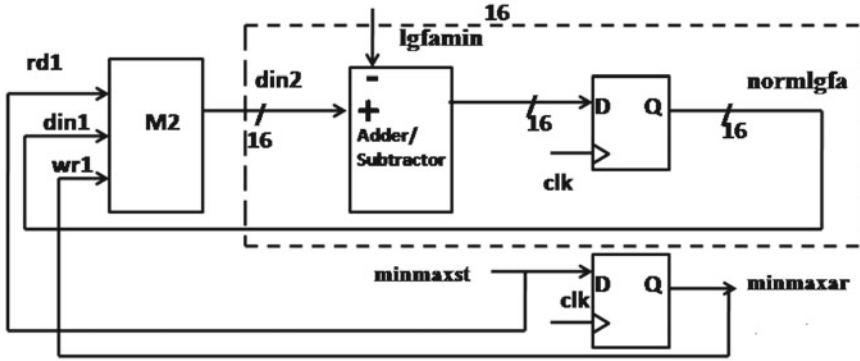
D:Datapath for Min-Max Normalization The min-max normalization in LGFA computation [11] involves a division of each LGFA offset minus the minimum LGFA angle in the LGFA matrix, by the absolute difference between maximum and minimum values in the matrix.

$$\bar{\alpha}_{\text{mm},c} = \frac{\alpha_c - \min(\alpha_c)}{|\max(\alpha_c) - \min(\alpha_c)|} \text{ where } \alpha_c - \min\alpha_c < |\max(\alpha_c) - \min(\alpha_c)| \quad (1)$$

In Eq. (1), $\bar{\alpha}_{\text{mm},c}$ is the min-max normalized LGFA, α_c is the computed LGFA at pixel position c and α_c the computed LGFA value normalized by 360° . This is incorporated into the architecture using the following steps. (1) Since the difference between maximum and minimum values of LGFAs cannot be more than 360° , the table of CORDIC’s reference angles is normalized by 360° . These angles are represented in fixed-point format Q1.14. Therefore, CORDIC computes LGFA directly normalized by 360° , i.e. α_c . Since the LGFA output of CORDIC, “lgfa” presents data sequentially, the register “lgfamin” for the minimum normalized LGFA $\min\alpha_c$, which is also updated sequentially by a simple compare and replace oper-



(a) Part of LGFA Datapath generating lgfamin



(b) The Min-Max Datapath

Fig. 6 LGFA computation datapath

ation : $lgfamin \leq lgfa$ when ($lgfa \leq lgfamin$). Then, at the end of all pixels, $lgfamin$ stores the $\min\alpha_c$. This computation datapath is shown on Fig. 6a. Note that this operation runs in parallel to process of storing of α_c in the memory M2. (2) After completion of all LGFA computation, each normalized LGFA α_c is read out sequentially from M2, subtracted by $lgfamin$, and then written back to the same address location in M2. The datapath for this step is shown in Fig. 6b. Naturally, M2 is implemented as a dual-port memory.

4 Result and Discussion

The proposed architecture is simulated and synthesized using Xilinx ISE 14.5 for Virtex-5 XC5VLX110T and code written in VHDL. The proposed generic hardware solution is scalable with respect to image dimensions and database size up to 128×128 . The number of bits for fixed-point numbers is chosen so that we do not lose out accuracy. Here, Q1.14 format of fixed-point arithmetic is chosen to represent data. For verification of the design, we simulated it in Xilinx ISE 14.5. The input 8×8 grayscale image is converted to COE file using MATLAB script and loaded in $4K \times 32$ single port BRAM memory using Xilinx ISE. Due to limitation of package, the

validation of computed result is done on-chip similar to built-in-self test (BIST). For that purpose, LGFA of input image is precomputed using the algorithm of Algorithm 1 in MATLAB script which we call ‘golden LGFA’, and converted to COE file. The output memory thus has addressed space twice that of input memory. The limitations of Xilinx System Generation the golden LGFA image’s COE file is loaded in the first half of the memory (between 0 and 4095). The second half of the memory is used for an intermediate result as well as final min-max normalized LGFA of the input image between 4096 and 8191. This assumes that the maximum image size is 128×128 . For this purpose, M2 memory is a dual-port memory. Finally, LGFA pixels are read from both first and second half of output memory and compared up to a tolerance value (0.065° in this case). A counter is maintained to have a record of matching LGFA pixels. The maximum difference between the golden and computed value for various image dimensions is compared in Table 1.

An example simulation result for the 8×8 image in this proposed hardware solution for LGface is shown in Fig. 7. The maximum difference between the golden and computed value is observed from the *diffd1d2* signal when the controller enters the *validate* state. This value is converted to a degree using the formula $(x/2^{14}) \times 360^\circ$, where x is the maximum value of *diffd1d2* signal under validate state. The *lgfamatch* signal is high when the pixels of golden lgfa data stored in the first half of memory M2 matches with computed lgfa pixel generated by proposed hardware and stored in the second half of M2 within a tolerance of 0.138° . An example simulation result for

Table 1 Simulated maximum difference of golden value computed by MATLAB and computed value of LGFA in proposed architecture in degree

Face image dimension	Maximum difference of golden ($^\circ$)
4×4 image	0.131
8×8 image	0.143
16×16	0.168
128×128	0.274



Fig. 7 Simulation result of 8×8 image for LGFA feature extraction

the 8×8 image in this proposed hardware solution for LGface is shown in Fig. 7. The maximum difference between the golden and computed value is observed from the *diffd1d2* signal when the controller enters the *validate* state. This value is converted to a degree using the formula $(x/2^{14}) \times 360^\circ$, where x is the maximum value of *diffd1d2* signal under *validate* state. The *lgfamatch* signal is high when the pixels of golden lgfa data stored in the first half of memory M2 matches with computed lgfa pixel generated by proposed hardware and stored in the second half of M2 within a tolerance of 0.138° . From synthesis results, the maximum operating frequency of proposed architecture is found to be 68.312 MHz. The runtime is 2.405 sec in computing the LGFA of 128×128 input image by MATLAB implementation when the proposed algorithm is executed on a computer with Intel(R) ‘Core(TM) i3-4030U CPU running at 1.93 GHz and 4 GB RAM. The total runtime of the proposed algorithm, when implemented in proposed hardware design is calculated from the start of lgfa state to the end of minmax calculation state and is 2.9193 ms. This shows that the proposed architecture is 2402. 1477 ms faster over MATLAB implementation. The device utilization is given in Table 2.

The power analysis for this LGFA hardware implementation is done for different image dimensions and summarized in Table 3. As can be observed from the chart, the total power is almost constant for different image dimensions. The maximum contribution to the power is due to leakage as we have selected a large device. The dynamic power is almost linear and constitutes 3.60% of the total power which is very small for the design. We performed experiments to evaluate the performance of the proposed architecture for the Extended Yale Face Database B (cropped) [25, 26].

Table 2 Device utilization of proposed architecture in virtex-5 XC5VLX110T

Logic utilization	Used	Available	Utilization (%)
No. of slice registers	455	69120	0.658
No. of slice LUTs	1724	69120	2
No. of fully Used LUT-FF pairs	259	1920	13
No. of bonded IOB	10	640	1
No. of block RAM FIFO	19	148	12
No. of DSP48ES	2	64	3

Table 3 Power analysis table

Image dimension	Total power (mW)	Dynamic power (mW)	Quiescent power(mW)
4×4	1.082	0.039	1.043
8×8	1.083	0.04	1.043
64×64	1.085	0.042	1.043
128×128	1.086	0.043	1.043

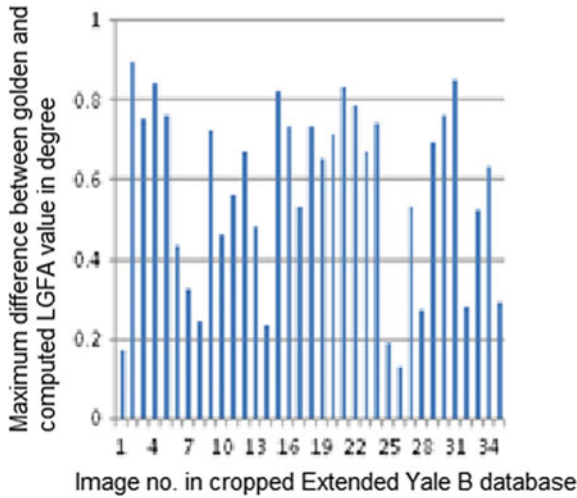


Fig. 8 Showing bar chart analysis on Extended Yale B (cropped)

Table 4 Comparison of the proposed LGFA architecture in different Xilinx platform

Image size	Platform	LUT	LUT-FF	Frequency (MHz)	Throughput/ No. of slices
128 × 128	Xilinx Spartan 3 Xc3s40	1962	489	40.038	14.707/1050 = 0.014
128 × 128	Xilinx Spartan 6 XC6SLX4	1692	277	74.742	7.874/465 = 0.016
128 × 128	Xilinx Virtex 5 XC5VLX110T	1724	259	85.390	6.897/455 = 0.015

The image size in Extended Yale Face Database B (cropped) is 192 × 168. All image data from Extended Yale Face Database B (cropped) are again manually cropped, and then resized to 128 × 128 images and then used in the experiment for this paper. The histogram chart for 35 images test images (128 × 128) Extended Yale B database is shown in Fig. 8. The result shows that the maximum error varies between 0.17° to 0.89° when query image is subjected to different illumination conditions. Similar experiments are performed on CMU-PIE database shows maximum error varies from 0.22° to 0.91° and CUFS database it varies from 0.19° to 0.94°.

The comparison of the proposed architecture on different Xilinx platform is given in Table 4. The comparison is made with respect to LUT, LUT-FF, frequency and throughput/number of slices. Throughput is generally used to compare the quality of competing architectures [27].

5 Conclusions

We have presented a novel VLSI architecture for generating LG-face . In architecture of [28], 64×64 image is used. In [29] is weighted modular principal component analysis (WMPCA) and face image dimension is not mentioned while in [30] image dimension is 32×32 . The face image dimension is 128×128 for the proposed architecture. It is implemented in FPGA showing that the hardware have superior performance compared to MATLAB implementation. The design can be used with a hardware implementation of classifier (like KNN) or the classifier can be invoked through the cloud via FPGA system to implement the face recognition using LGFA extraction method.

References

1. W. Wójcik, K. Gromaszek, M. Junisbekov, Face recognition: Issues, methods and alternative applications. *Face Recognition-Semisupervised Classification, Subspace Projection and Evaluation Methods* (2016)
2. R.C. Gonzalez, R.E. Woods, S.L. Eddins, *Digital Image Processing Using MATLAB* (Pearson Educ, India, 2004)
3. M. Savvides, B.V. Kumar, Illumination normalization using logarithm transforms for face authentication, in *International Conference on Audio-and Video-Based Biometric Person Authentication* (Springer, 2003), pp. 549–556
4. R. Basri, D.W. Jacobs, Lambertian reflectance and linear subspaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(2), 218–233 (2003)
5. P.N. Belhumeur, D.J. Kriegman, What is the set of images of an object under all possible illumination conditions? *Int. J. Comput. Vis.* **28**(3), 245–260 (1998)
6. Y. Pang, Y. Yuan, X. Li, Gabor-based region covariance matrices for face recognition. *IEEE Trans. Circuits Syst. video Technol.* **18**(7), 989–993 (2008)
7. T. Ahonen, A. Hadid, M. Pietikainen, Face description with local binary patterns: application to face recognition. *IEEE Trans. Patt. Anal Mach. Intell.* **28**(12), 2037–2041 (2006)
8. M.K. Bhowmik, D. Bhattacharjee, D.K. Basu, M. Nasipuri, Eye region-based fusion technique of thermal and dark visual images for human face recognition. *Opt. Eng.* **51**(7), 077,205 (2012)
9. S.Z. Li, R. Chu, S. Liao, L. Zhang, Illumination invariant face recognition using near-infrared images. *IEEE Trans. Patt. Anal. Mach. Intell.* **29**(4), 627–639 (2007)
10. J.Y. Zhu, W.S. Zheng, J.H. Lai, S.Z. Li, Matching nir face to vis face using transduction. *IEEE Trans. Inf. Forensics Secur.* **9**(3), 501–514 (2014)
11. H. Roy, D. Bhattacharjee, Local-gravity-face (lg-face) for illumination-invariant and heterogeneous face recognition. *IEEE Trans. Inf. Forensics Secur.* **11**(7), 1412–1424 (2016)
12. B.F. Klare, A.K. Jain, Heterogeneous face recognition using kernel prototype similarities. *IEEE Trans. Patt. Anal. Mach. Intell.* **35**(6), 1410–1422 (2012)
13. Z. Lei, S. Liao, A.K. Jain, S.Z. Li, Coupled discriminant analysis for heterogeneous face recognition. *IEEE Trans. Inf. Forensics Secur.* **7**(6), 1707–1716 (2012)
14. H. Roy, D. Bhattacharjee, Local-gravity-face (lg-face) for illumination-invariant and heterogeneous face recognition. *IEEE Trans. Inf. Forensics Secur.* **11**(7), 1412–1424 (2016). <https://doi.org/10.1109/TIFS.2016.2530043>
15. J.M. Gilbert, W. Yang, A real-time face recognition system using custom VLSI hardware. *IEEE* (1993)

16. G. Theoretical, H. Mahale, A. Goel, S. Nandy, S. Bhattacharya, R. Narayan, Hardware solution for real-time face recognition, in *2015 28th International Conference on VLSI Design (IEEE, 2015)*, pp. 81–86
17. J. Matai, A. Irturk, R. Kastner, Design and implementation of an fpga-based real-time face recognition system, in *2011 IEEE 19th Annual International Symposium on Field-Programmable Custom Computing Machines (IEEE, 2011)*, pp. 97–100
18. F. Yang, M. Paindavoine, Implementation of an rbf neural network on embedded systems: real-time face tracking and identity verification. *IEEE Trans. Neural Networks* **14**(5), 1162–1175 (2003)
19. S. Kang, J. Lee, C. Kim, H. Yoo, B-face: 0.2 mw cnn-based face recognition processor with face alignment for mobile user identification, in *2018 IEEE Symposium on VLSI Circuits*, (2018), pp 137–138. <https://doi.org/10.1109/VLSIC.2018.8502266>
20. M. Madhivhanan, R. Ravi, Vlsi implementation multimodal biometric recognition system using improved pulse coupled neural network for digital nano-cmos technology. *J. Comput. Theor. Nanosci.* **14**(11), 5657–5662 (2017)
21. J.E. Volder, The cordic trigonometric computing technique. *IRE Trans. Electron. Comput.* **EC-8**(3), 330–334 (1959)
22. J.S. Walther, The story of unified cordic. *J. VLSI Sig. Proc. Syst. Sig. Image video Technol.* **25**(2), 107–112 (2000)
23. P.K. Meher, J. Valls, T.B. Juang, K. Sridharan, K. Maharatna, 50 years of cordic: algorithms, architectures, and applications. *IEEE Trans. Circ. Syst. I: Regul. Pap.* **56**(9), 1893–1907 (2009)
24. R. Andraka, A survey of cordic algorithms for fpga based computers, in *Proceedings of the 1998 ACM/SIGDA Sixth International Symposium on Field Programmable Gate Arrays* (1998), pp. 191–200
25. A. Georghiades, P. Belhumeur, D. Kriegman, From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(6), 643–660 (2001)
26. K.C. Lee, J. Ho, D.J. Kriegman, Acquiring linear subspaces for face recognition under variable lighting. *IEEE Trans. Patt. Anal. Mach. Intell.* **27**(5), 684–698 (2005)
27. N. Cortez-Duarte, F. Rodríguez-Henríquez, J.L. Beuchat, E. Okamoto, A pipelined Karatsuba-Ofman multiplier over $gf(397)$ amenable for pairing computation. *IACR Cryptol ePrint Arch* **2008**, 127 (2008)
28. R. Gottumukkal, H.T. Ngo, V.K. Asari, Multi-lane architecture for eigenface based real-time face recognition. *Microprocess. Microsyst.* **30**(4), 216–224 (2006)
29. A.P. Kumar, V. Kamakoti, S. Das, System-on-programmable-chip implementation for on-line face recognition. *Patt. Recog. Lett.* **28**(3), 342–349 (2007)
30. N. Sudha, A. Mohan, P.K. Meher, A self-configurable systolic architecture for face recognition system based on principal component neural network. *IEEE Trans. Circuits Syst. Video Technol.* **21**(8), 1071–1084 (2011)

Enhanced Performances of SAC-OCDMA System Operating with Different Codes



Walid Sahraoui, Hakim Aoudia, Angela Amphawan, Smail Berrah, Youcef Belkhier, and Rabindra Nath Shaw

Abstract Herein, spectral amplitude coding-optical code division multiple access (SAC-OCDMA) technology is investigated and simulated. Thus, a comparative analysis is done between the random diagonal (RD), modified double weight (MDW) and enhanced double weight (EDW) codes, aiming to obtain results with better performances. The suggested system that exploits lasers as transmitter shows significantly perfect performances as compared to the bit error rate (BER), eye's diagram, and Q factor of others standard SAC-OCDMA.

Keywords SAC-OCDMA · RD · EDW · MDW · BER · Optical fiber

1 Introduction

Optical fiber communication is the key technology behind the endless growth of Internet. However, such technology suffers from some phenomena like the linear and nonlinear effects (attenuation, dispersion, Kerr effect, ..., etc.). These effects degrade the quality of signal and increase noises during the transmission. Therefore,

W. Sahraoui (✉) · H. Aoudia · S. Berrah

Laboratoire de Maitrise des Energies Renouvelables (LMER), Faculté de Technologie, Université de Bejaia, 06000 Bejaia, Algeria

W. Sahraoui · A. Amphawan

Lightwave Research Laboratory, School of Engineering and Technology, Sunway University, Petaling Jaya, Selangor, Malaysia

A. Amphawan

Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, USA

Y. Belkhier

Laboratoire de Technologie Industrielle et de l'Information (LTII), Faculté de Technologie, Université de Bejaia, 06000 Bejaia, Algeria

R. N. Shaw

Department of Electronics and Communication Engineering, Galgotias University, Greater Noida, India

the optical fiber communication requires multiplexing techniques that support a large number of simultaneous users demanding voice, video, and data services [1]. Also, reduce the linear and nonlinear phenomena's influence.

The OCDMA allows multiple users by assigning different optical codes (OCs) to share the same transmission medium to different users [2, 3]. Nowadays, OCDMA cannot be considered as a competitor of MDW, but it can be considered in optical access networks [4]. In the other side, the OCDMA systems suffer from different noises like the phase-induced intensity noise (PIIN), the dark current, add to that the multi-access interference (MAI) [5] who represents a complicated problem. In order to eliminate the MAI effect, various encoding techniques are proposed by many researchers, such as spectral phase coding but it was very difficult to preserve in fiber. Therefore, the spectral amplitude coding was proposed [6].

There are many different codes which have been considered for the SAC-OCDMA networks. In particular, EDW code, Khazani–Syed (KS) code, modified frequency hopping (MFH) code, RD code, MDW code, and ZFD code [5–7].

In this paper, a system based on the SAC-OCDMA technique was created. The RD, MDW, and EDW codes were applied. The aim of this work is to obtain enhanced performances, where a comparative analysis will be done between these codes.

2 Construction of Codes

2.1 The Random Diagonal

The random diagonal (RD) code uses unipolar code (0, 1) sequences. It is denoted by (N, w, λ) which N represents the length of the code, w represents the weight of the code, and λ denotes cross-correlation of the in-phase that has two proprieties. The first propriety represents the cross correlation of the zero that will minimize the λ and reduce the PIIN, and the second propriety is the in-data level no cross correlation [6, 8].

Data segment

This part is denoted by the matrix $(K * K)$, where K denotes the user's number and $w = 1$ to keep the zero cross-correlation at $\lambda = 0$. For $K = 3$, the matrix $(K * K)$ can be given as:

$$[Y_1] = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad (1)$$

Code segment

This second part will be represented by $[Y_2]$ matrix, which is the combination of two matrices (weight and basic matrix). The weight matrix $[M] = [M_1 M_2 M_3 \dots M_i]$ is

responsible for increasing the value of ‘ w ’ [6]. Where $i = w - 3$ and for $w = 4$, the weight matrix is given by $[M] = [M_1]$, and it is represented as follows:

$$[M] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{2}$$

The basic matrix $[B]$ can be expressed as:

$$[B] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \tag{3}$$

To increase the number of users or the code length, just repeat each range of the $[M]$ and $[B]$ matrices, and to get the $[Y_2]$ matrix, we will just combine the $[B]$ and $[M]$ matrices.

$$[Y_2] = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix} \tag{4}$$

The combination of $[Y_1]$ and $[Y_2]$ gives the total code sequence $[Z]$ with $N = 8$ and $K = 3$.

$$[Z] = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \tag{5}$$

2.2 MDW Code

The modified double weight (MDW) code is developed based on the double weight (DW) code of ($w = 2$) that has a large number of weights. This code is denoted by (N, w, λ) , and it has an ideal cross correlation ($\lambda = 1$) [9, 10]. MDW code is denoted by the matrix $[K * N]$ where K represents the user’s number and N represents the code of the length. For ($w = 4$), it is denoted by $(N, 4, 1)$ where the length of the code ‘ N ’ is related to the user’s number of ‘ K ’ through:

$$N = 3K + \frac{3}{8} \left[\sin \left(\frac{K\pi}{3} \right) \right]^2 \tag{6}$$

The basic MDW code of three users is given as follows [9]:

$$[K \times N] = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (7)$$

2.3 EDW Code

Enhanced DW code weight can be any odd number greater than one [10]. ESW code is denoted by (N, w, λ) , as in the RD and MDW codes. It has an ideal cross-correlation ($\lambda = 1$) where the relation between the number of users and the length code is given by [11]:

$$N = 2K + \frac{4}{3} \left[\sin\left(\frac{K\pi}{3}\right) \right]^2 + \frac{8}{3} \left[\sin\left(\frac{(K+1)\pi}{3}\right) \right]^2 + \frac{4}{3} \left[\sin\left(\frac{(K+2)\pi}{3}\right) \right]^2 \quad (8)$$

The EDW code is represented by the matrix $[K * N]$ given below [11]:

$$[K * N] = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{bmatrix} \quad (9)$$

3 System Description

The SAC-OCDMA system architecture employs different codes (RD, EDW and MDW) consisting of three active users, where the performances of these codes are tested, using Optisystem7.0 software [12]. The data signal of each user is generated from the pseudo random bit generator before using the NRZ coding. Then, the data signal is modulated with different codes using the external modulation (Fig. 1). In this system, the transmitted power is based on laser sources representing the sending wavelengths with 1.2 nm spacing between two adjacent wavelengths, to represent the code words (Tables 1, 2 and 3). After the combination of signals, one signal travels through the single mode fiber (SMF) and broadcasts to the receivers. Table 4 shows different used parameters during the simulations.

In the receiver side (Fig. 2), the signal is decoded using the fiber Bragg grating (FBG), where the same wavelengths in the encoding part for each user are employed. Then, the photo detector (PIN) is used to convert the optical signal to an electrical signal. The resultant signal is filtered by fourth order Bessel low-pass filter, which

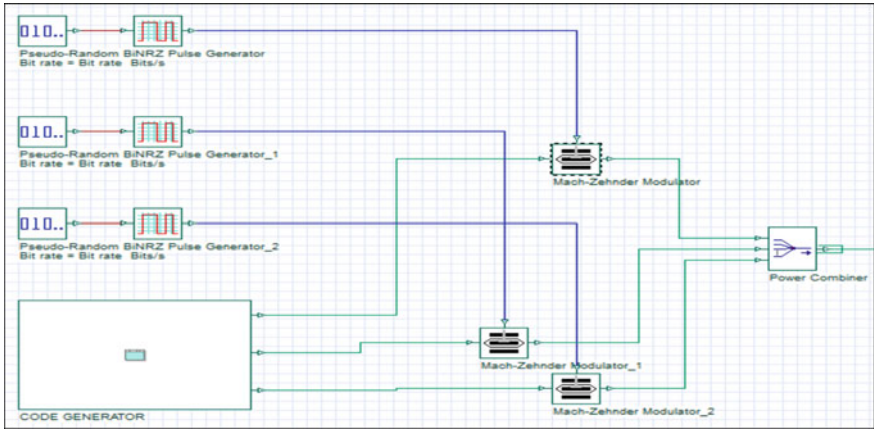


Fig. 1 Transmitter block of SAC-OCDMA system

Table 1 Wavelengths using in the RD code

Users	λ_1 (nm)	λ_2 (nm)	λ_3 (nm)	λ_4 (nm)
User.1	1552.4	1553.6	1554.8	1557.2
User.2	1551.2	1554.8	1556	1558.4
User.3	1550	1553.6	1556	1557.2

Table 2 Wavelengths using in the MDW code

Users	λ_1 (nm)	λ_2 (nm)	λ_3 (nm)	λ_4 (nm)
User.1	1554.8	1556	1558.4	1559.6
User.2	1551.2	1552.4	1557.2	1558.0.4
User.3	1550	1551.2	1553.6	1554.8

Table 3 Wavelengths using in the EDW code

Users	λ_1 (nm)	λ_2 (nm)	λ_3 (nm)
User.1	1552.4	1553.6	1556
User.2	1551.2	1554.8	1556
User.3	1550	1551.2	1553.6

is used to reject noise and interference components that lie outside the information signal spectrum [11]. Finally, each signal is analyzed by the BER analyzer.

Table 4 System parameters

Parameters	Value
Number of users	3
Receiver filter bandwidth	$1.8e-023$ W/Hz
Thermal noise coefficient	$0.75 \times$ bit rate
Signal format	NRZ
Signal data	1.2 nm
Attenuation	0.2 dB/km
External modulator extinction ratio	30 dB
Dispersion	16.75 ps/nm/km
Chip spectral width	128 PN sequence
Dark current	10 nA

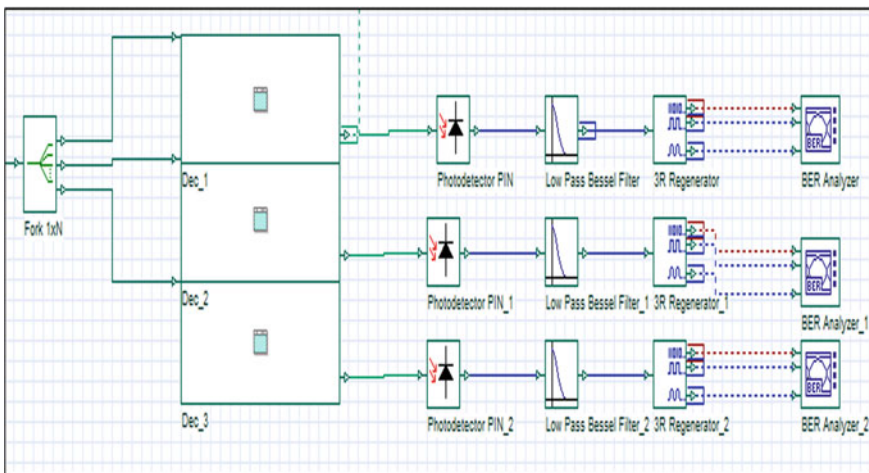


Fig. 2 SAC-OCDMA system diagram

4 Results and Discussion

In this paper, SAC-OCDMA system containing three users was created. The performance parameters like the bit error rate (BER), Q factor, and the eye diagram were tested by exploiting different codes (RD, EDW and MDW). A comparative analysis between these codes was realized. The performance parameters were tested according some criteria such as the fiber length and the source power [12–16].

Figures 3 and 4, respectively, show the BER and Q factor of three codes over the fiber length.

At each code simulation, the source power is fixed at 0 dBm and the data rate at 622 Mb/s. Herein, with fiber length, the BER value of all codes increases. The

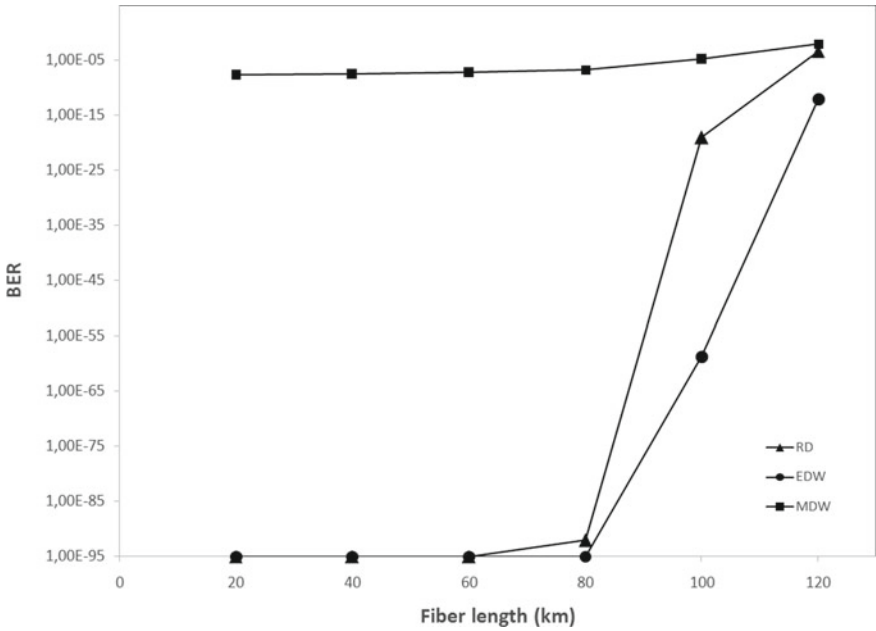


Fig. 3 BER versus fiber length at data rate of 622 Mbps

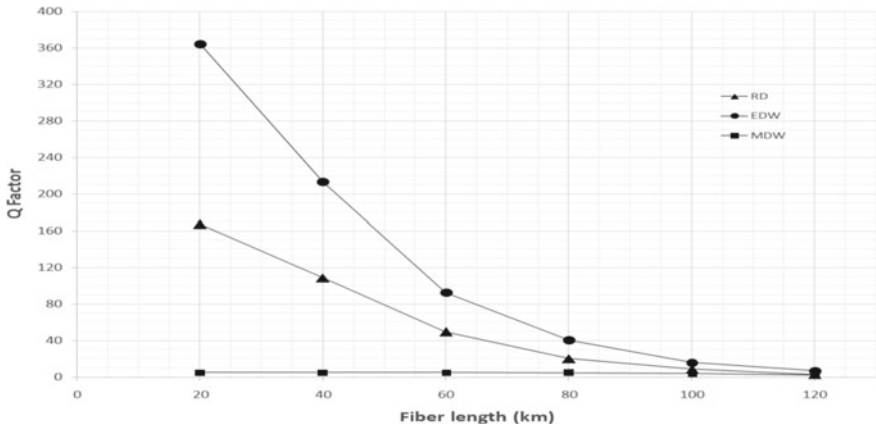


Fig. 4 Q factor versus fiber length at data rate of 622 Mbps

EDW code gives better results than the other codes with a perfect BER value until a distance of 80 km, and it is around 10^{-13} at a length of 120 km. Even the RD code has good results with perfect BER value until a length of 60 km and around 10^{-4} at 120 km. Similarly, the Q factor was obtained, except that it decreases with the increasing of the fiber length. In [6], the AGW, FGB, and thin film filter have been

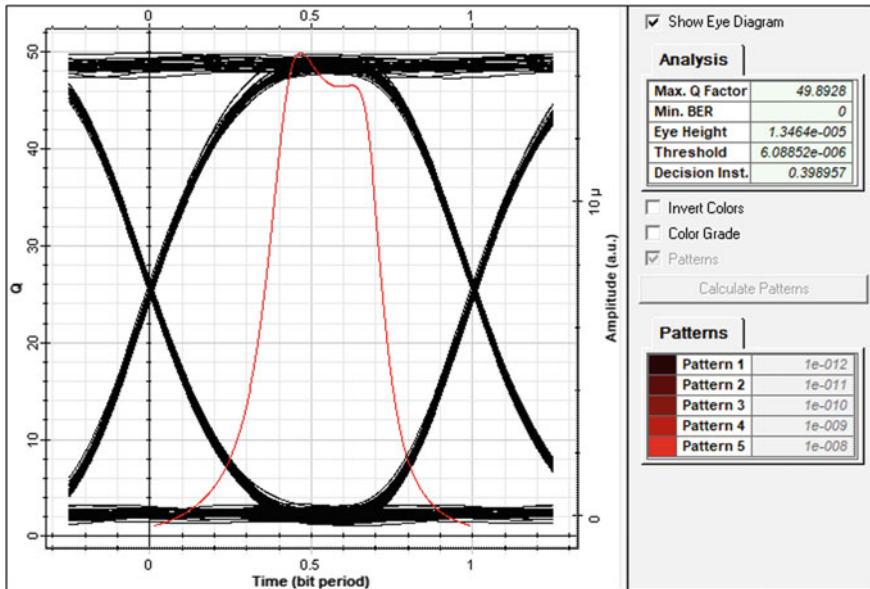


Fig. 5 RDcode eye's diagram at 60 km

tested as decoders. They have used the RD code. For the FBG, the BER was around 10^{-15} at length of 40 km.

The eye diagrams represented in Figs. 5, 6, and 7 clearly show that EDW code is better. Even RD code has too large eye opening. In this test, MDW always does not give good results comparing to the other two codes.

The second test is based on the source's power, where the fiber length of 80 km and a data rate of 622 Mbps are set as link parameters.

Figure 8 shows the BER parameter of different power values. The results that were found show the decreasing of value, where a perfect result is reached in both RD and EDW codes from a source's power of 5 dBm. The same thing about the Q factor is represented in Fig. 9 except that the Q factor values increase whenever the increasing of the power values of source. This increasing requires more cost and high energy which are considered as disadvantages. Therefore, it is preferable to use a 0 dBm source's power as an optimal solution.

The eye diagrams of each code at a source's power of 5 dBm are represented in Figs. 10, 11, and 12. As in all simulations, the best results are given by the EDW code regardless of the tested parameter even the RD code has good results. It is clearly to see that those both codes have a larger opening of the eye's diagram than the MDW code.

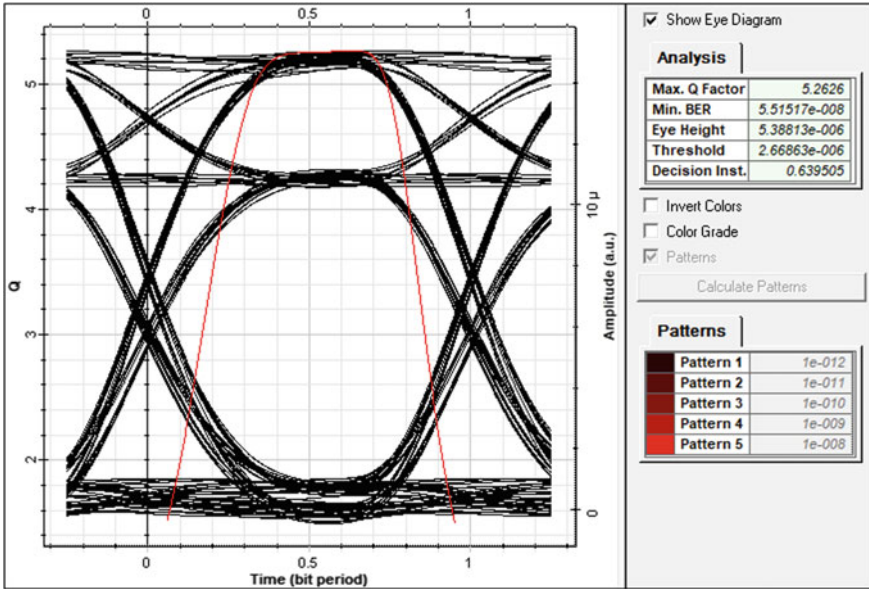


Fig. 6 MDW code eye's diagram at 60 km

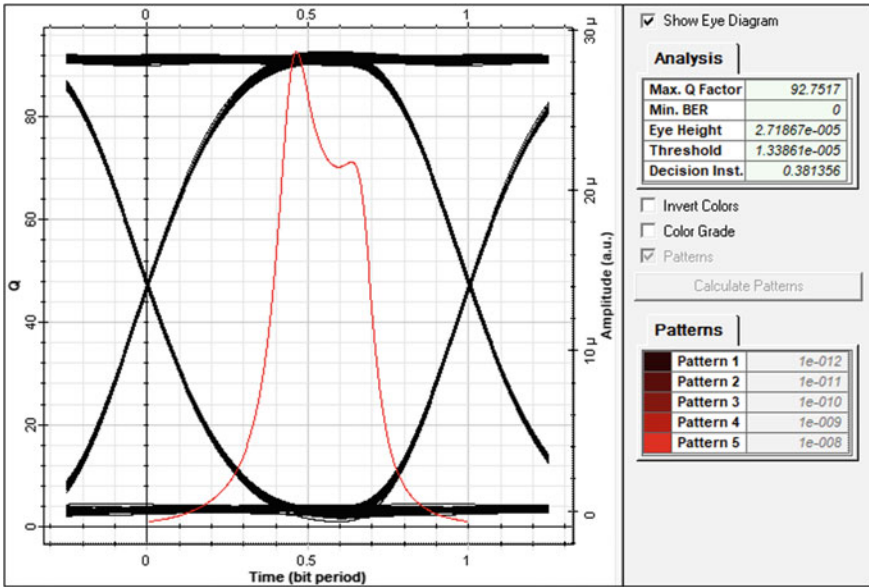


Fig. 7 EDW code eye's diagram at 60 km

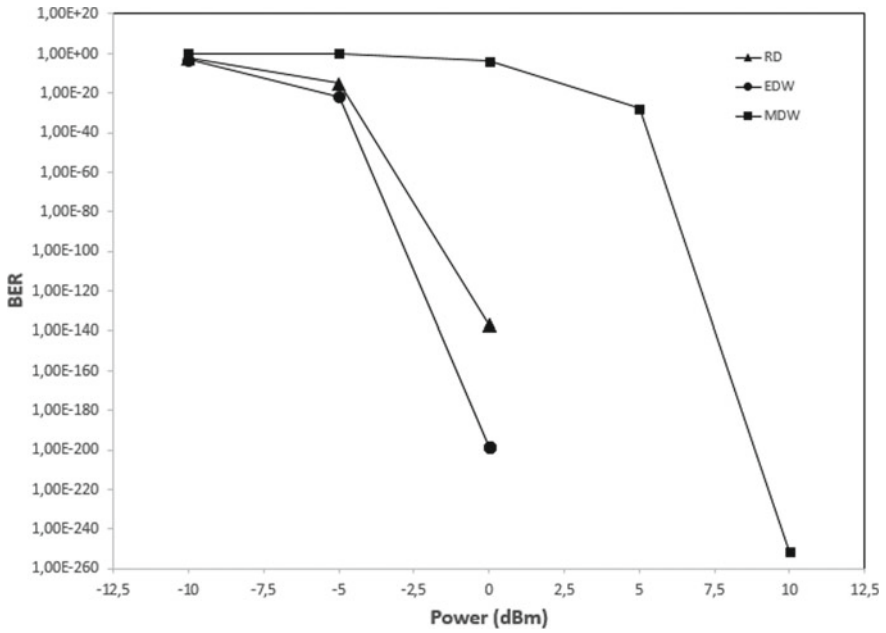


Fig. 8 BER versus the power of source

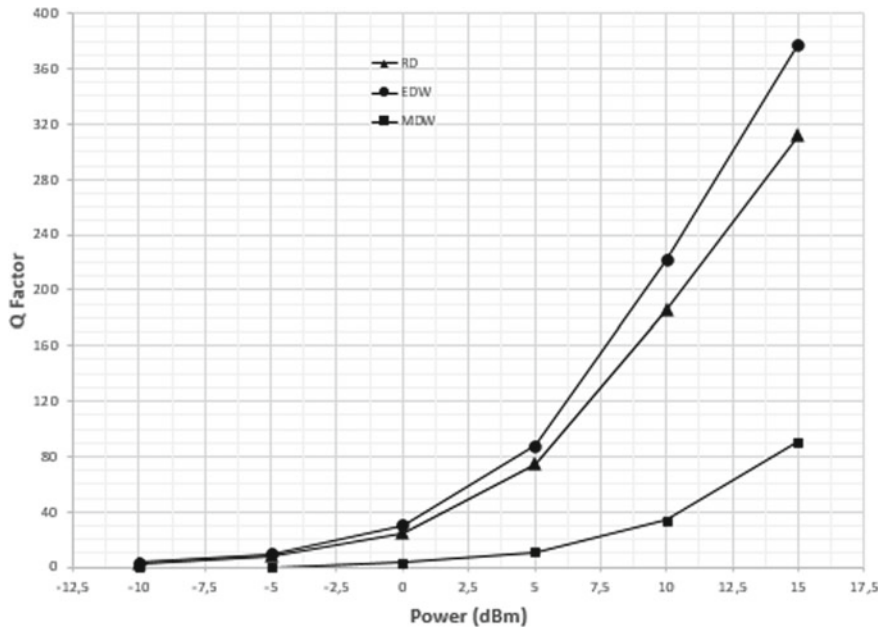


Fig. 9 Q factor versus the power of source

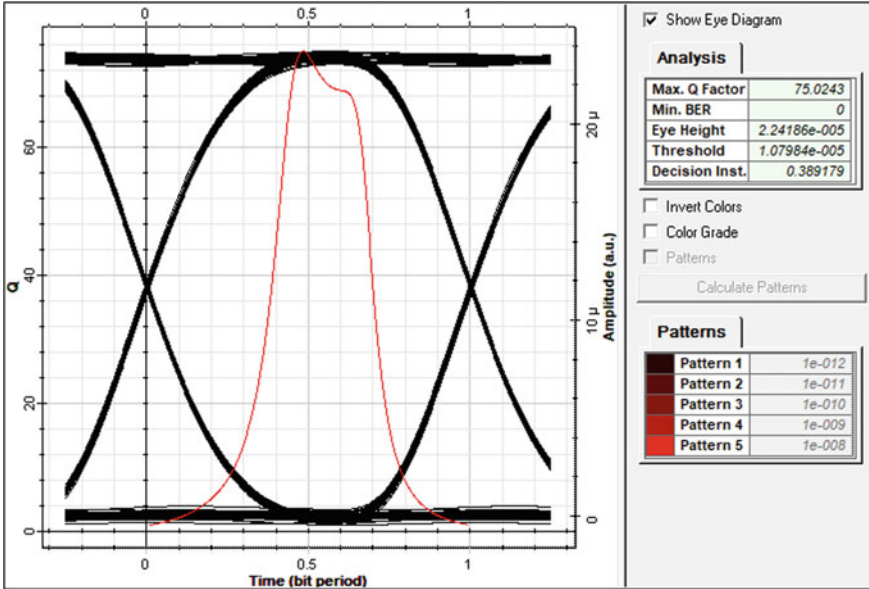


Fig. 10 Eye's diagram of RD code at 5 dBm

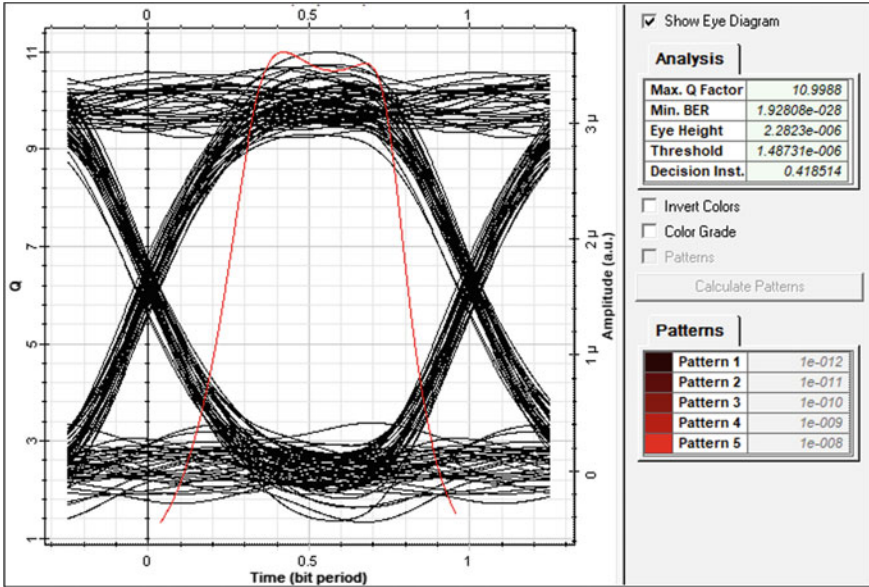


Fig. 11 Eye's diagram of MDW code at 5 dBm

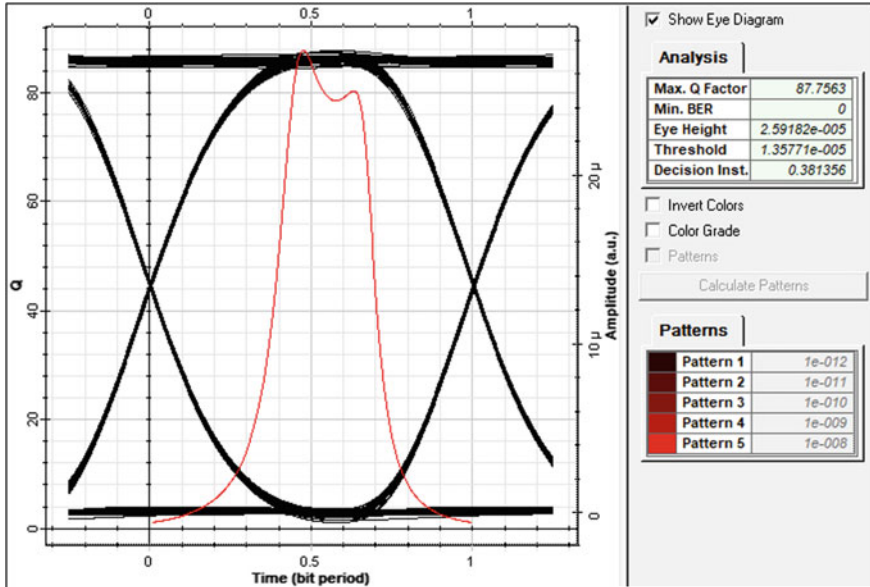


Fig. 12 Eye's diagram of EDW code at 5 dBm

5 Conclusion

In this paper, an SAC-OCDMA system containing three users was created, where the following performance parameters like the bit error rate (BER), Q factor, and the eye's diagram were tested by exploiting different codes which are the RD, EDW, and MDW. This study had the object of reach enhancement performances of the optical link than the others standard SAC-OCDMA. Therefore, a comparative analysis between these codes has been done. Based on this analysis, the EDW had always the better results than the other codes, despite it exploited the minimum number of wavelengths. The performance parameters tests were according to different criteria like the fiber length and the source's power.

References

1. X. Chen, D. Chen, Z. Wang, Performance improvement of bandwidth-limited coherent OCDMA system. *Photonic Netw. Commun.* **16**(2), 149–154 (2008)
2. H. Yan, K. Qiu, Y. Ling, Performance analysis of a hybrid OCDMA/WDM system. *Front. Optoelectron. China* **2**(1), 15–20 (2009)
3. A.O. Aldhaibani, T.A. Rahman, S.A. Aljunid, M.N. Hindia, E.B. Hanafi, A new model to enhance the QoS of spectral amplitude coding-optical code division multiple access system with OFDM technique. *Opt. Quantum Electron.* **48**(10), 1–12 (2016)

4. X. Chen, X. Yuan, Investigation of superstructured fiber Bragg grating-based OCDMA system. *Photonic Netw. Commun.* **20**(3), 242–246 (2010)
5. T. Abd, S. Aljunid, H. Fadhil, R. Ahmad, M. Junita, Enhancement of performance of a hybrid SAC-OCDMA system using dynamic cyclic shift code. *Ukr. J. Phys. Opt.* **13**(1), 12 (2012)
6. H. Adhil, S. Aljunid, B. Ahmed, Performance of OCDMA systems using random diagonal code for different decoders architecture schemes. *Int. Arab J. Inf. Technol.* **7**(1), 1–5 (2010)
7. W. Sahraoui, H. Aoudia, S. Berrah, A. Amphawan, R. Naoum, Performances analysis of novel proposed code for SAC-OCDMA system. *J. Opt. Commun.* 1(ahead-of-print) (2020)
8. M. Saxena, R.N. Shaw, J.K. Verma, A novel hash-based mutual RFID tag authentication protocol, in *Advances in Intelligent Systems and Computing*, vol. 847 (Springer, Berlin, 2019), pp. 1–12. https://doi.org/10.1007/978-981-13-2254-9_1
9. I. Das, R.N. Shaw, S. Das, Analysis of energy consumption of energy models in wireless sensor networks, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh, vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_57
10. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh, vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
11. I. Das, R.N. Shaw, S. Das, Location-based and multipath routing performance analysis for energy consumption in wireless sensor networks, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh, vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_59
12. I. Das, R.N. Shaw, S. Das, Analysis of energy consumption in dynamic mobile ad hoc networks, in *Advances in Intelligent Systems and Computing*, vol. 1049 (Springer, Berlin, 2020), pp. 235–243. https://doi.org/10.1007/978-981-15-0132-6_15
13. F. Su, H. Jin, Research of code construction for OCDMA system. *Adv. Mech. Electron. Eng.* **178**, 265–270 (2013)
14. S.G. Abdulqader, S.A. Aljunid, H.M.R. Al-khafaji, H.A. Fadhil, Enhanced performance of SAC-OCDMA system based on SPD detection utilizing EDFA for access networks. *J. Commun.* **9**(2), 99–106 (2014)
15. F.N. Hasoon, M.H. Al-Mansoori, S. Shaari, Different detection schemes using enhanced double weight code for OCDMA systems, in *Systems IAENG Transactions on Engineering Technologies IAENG Transactions on Engineering Technologies*, vol. 170 (2013), pp. 287–299
16. S.A. Abd E Mottaleb, H.A. Fayed, A. Abd El Aziz, M.H. Aly, SAC-OCDMA system using different detection techniques. *IOSR J. Electron. Commun. Eng.* **9**(2), 55–60 (2014)

Phasor Estimation of Power Signals in a Smart Grid Environment Using Sigmoid Adaptive Filter



Soumili Saha, Prateek Bose, and Sarita Nanda

Abstract A smart grid is an electric grid that provides two-way communication between production and consumption. The integration of different sources of power generation makes it vulnerable to disturbances between the supply and demand; therefore, control and tracking of parameters are necessary for the smart grid system. To overcome the amplitude and phase distortion along with harmonic disturbances, this paper uses phasor estimation techniques. In this paper, we focus on some algorithms based on the modified least mean mixed norm (LMMN) for the same and compare the results for a test case to solve the extensive and exclusive issues of smart grid phasor estimation.

Keywords Adaptive algorithm · Phasor estimation · Least mean mixed norm (LMMN) · Sigmoid least mean mixed norm (SLMMN)

1 Introduction

To meet the ever-changing and rising demands of the world, the necessity of monitoring and control over the electric power system is becoming progressively important. The smart grid provides a two-way route for energy and information exchange; this makes the system more efficient and reliable. The smart grid enables different power sources to be integrated such as solar and wind energy production, thus paving the way to renewable resource utilization [1]. But these power sources are not governed; therefore, the addition of these different sources of energy to the existing system brings in the challenge of handling information and distribution. Also, this brings the need to overcome the amplitude and phase distortion along with the elimination of harmonic disturbances that occur [2]. This is done by using several adaptive algorithms for the phasor estimation of the power signal [3]. This paper introduces a design for a sigmoid adaptive filter for phasor measurement in the smart grid. Therefore, agglomerating different weight signals updation for phasor estimation.

S. Saha · P. Bose · S. Nanda (✉)
School of Electronics, KIIT University, Bhubaneswar, Odisha, India
e-mail: snandafet@kiit.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_36

487

Section 2 concisely summarizes phasor measurement using sigmoid-based algorithm, while Sect. 3 provides simulation for non-stationary signals. Section 4 deals with some conclusion point at the end.

2 Sigmoid-Based Algorithm for Phasor Measurement

The desired analog voltage [2] is represented as the sum of the fundamental signal and its odd harmonic signals as:

$$v(k) = \sum_{h=1}^H A_h(k) \cdot \sin(2\pi hf(k)dt + \phi_h(k)) \tag{1}$$

where f is the fundamental frequency, $A_h(k)$ is the time-varying amplitude of the h th harmonic component at the k th sampling instant, $\phi_h(k)$ is the phase angle of the h th harmonic component at the k th sampling instant, dt is the sampling time, and h refers to the order of harmonic [4].

$$v(k) = A_1(k) \cdot \cos(\phi_1(k)) \cdot \sin(2\pi f(k)dt) + \dots + A_H(k) \cdot \sin(\phi_H(k)) \cdot \cos(2\pi Hf(k)dt) \tag{2}$$

can be further expanded as:

$$v(k) = W_{a1} \sin(2\pi f(k)dt) + W_{b1} \cos(2\pi f(k)dt) + W_{a3} \sin(2\pi 3f(k)dt) + W_{b3} \cos(2\pi 3f(k)dt) + \dots + W_{ah} \sin(2\pi hf(k)dt) + W_{bh} \cos(2\pi hf(k)dt) \tag{3}$$

where the weights W_{ah} and W_{bh} are given as $W_{ah} = A_h(k) \cos(\phi_h(k))$, and $W_{bh} = A_h(k) \sin(\phi_h(k))$.

The weight vector is expanded as:

$$\widehat{W}(k) = [\widehat{W}(k)_{a1} \widehat{W}(k)_{b1} \dots \widehat{W}(k)_{ah} \widehat{W}(k)_{bh}] \tag{4}$$

The input vector can be expanded as:

$$\widehat{U}(k) = [\sin(2\pi f(k)dt) \cos(2\pi f(k)dt), \dots, \sin(2\pi hf(k)dt) \cos(2\pi hf(k)dt)] \tag{5}$$

The ADALINEF vector is obtained by multiplying the input coefficient vector and weight vector. It is represented as:

$$\hat{v}(k) = \hat{U}(k-1).W^T(k) \quad (6)$$

The estimated error determined is represented as:

$$\varepsilon(k) = v(k) - \hat{v}(k) \quad (7)$$

$$\varepsilon(k) = v(k) - \hat{U}(k-1).W^T(k) \quad (8)$$

The weight signal is a function of a variable k .

LMMN algorithm mortify the performance of the phasor estimation due to which a sigmoid-based LMMN algorithm is used to overcome the sensitivities of LMMN algorithm. Therefore, SLMMN is proposed for determining the optimized phasor properties like amplitude and phase estimation [5–7].

The weight vector can be updated as given in the following section.

2.1 Least Mean Mixed Norm

Least mean squares (LMS) algorithm use estimation of gradient vector for adaptive filtration by the modification of the mathematical function [8].

Combining the cost functions of LMS and least mean fourth (LMF) algorithms, the least mean mixed norm algorithm is conferred as:

$$I_{\text{LMMN}}(k) = \frac{\gamma}{2} E\{\varepsilon^2(k)\} + \frac{1-\gamma}{4} E\{\varepsilon^4(k)\} \quad (9)$$

$$\begin{aligned} W(k+1) &= W(k) - \mu \hat{\nabla} W(k) I_{\text{LMMN}}(k) \\ &= W(k) + \varepsilon(k) \{Y + (1-\gamma)\varepsilon^2(k)\} U(k) \end{aligned} \quad (10)$$

where γ is the mixing parameter and μ is the step size.

$$W(k) = W(k) + \mu.\varepsilon.(L + ((1-L).\varepsilon^2)).\hat{U}(k) \quad (11)$$

where μ and L are constants.

2.2 Sigmoid-Based Least Mean Mixed Norm

The main purpose to use sigmoid function is that it is limited between the range of 0–1 and is differentiable. Representation of sigmoid function in terms of cost function is represented as:

$$s(k) = \text{sgm}[\beta \cdot I_{\text{LMMN}}(k)] = \frac{1}{e^{-\beta I_{\text{LMMN}}(k)} \left[\frac{1}{e^{-\beta I_{\text{LMMN}}(k)} + 1} \right]} \tag{12}$$

where $s(k)$ is the sigmoid cost function and β is the parameter of steepness.

The sigmoid cost function of the modified LMMN algorithm is used to overcome impulse interference of external environment is given by:

$$I_{\text{SLMMN}}(k) = \frac{1}{\beta} S(k) = \frac{1}{\beta} \frac{1}{1 + e^{-\beta I_{\text{LMMN}}(k)}} \tag{13}$$

$$W(k + 1) = W(k) - \mu \hat{\nabla}_{W(k)} I_{\text{SLMMN}}(k) \tag{14}$$

The fundamental and harmonic functions are formulated in terms of weight function. It can be obtained from:

$$\begin{aligned} A_1(k) &= \sqrt{W^2(1) + W^2(2)} \\ A_3(k) &= \sqrt{W^2(3) + W^2(4)} \\ A_5(k) &= \sqrt{W^2(5) + W^2(6)} \dots \end{aligned} \tag{15}$$

The phase values are obtained from (Fig. 1):

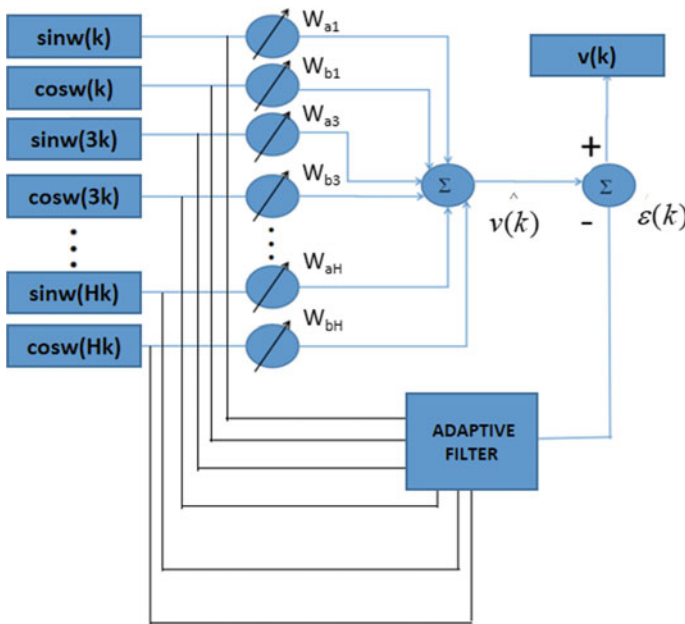


Fig. 1 Generalized structure of adaptive filter

$$\begin{aligned}
 \phi_1(k) &= a \tan(W(2)/W(1)) \\
 \phi_3(k) &= a \tan(W(4)/W(3)) \\
 \phi_5(k) &= a \tan(W(6)/W(5)) \dots
 \end{aligned}
 \tag{16}$$

3 Simulation and Results

3.1 Non-stationary Signal Analysis

The test signal can be represented as:

$$\begin{aligned}
 v(k) &= A_1(k) \cdot \sin(\omega(k)k + \phi_1(k)) + A_3(k) \cdot \sin(3\omega(k)k + \phi_3(k)) \\
 &\quad + A_5(k) \cdot \sin(5\omega(k)k + \phi_5(k))
 \end{aligned}
 \tag{17}$$

where $A_1, A_3,$ and A_5 are the amplitudes and $\phi_1, \phi_3,$ and ϕ_5 are the phase angles of the harmonic components. The different base parameters of (17) are given in Table 1. The estimation results using LMMN and SLMMN are presented as follows:

Figures 2, 3, and 4 represent the phasor estimation for the test signal (17). From here, it is found that SLMMN provides the most accurate result of among these two algorithms, namely LMMN and SLMMN, and Table 2 represents error for the components under different SNR conditions.

Table 1 Variation in amplitude and phase components of test signal

Samples	A_1	A_3	A_5	ϕ_1	ϕ_3	ϕ_5
0–201	1.2	0.5	0.3	−0.3927	−1.0472	−0.4488
201–448	0.96	0.5	0.3	−0.3927	−1.0472	−0.4488
448–601	1.44	0.5	0.3	−0.3927	−1.0472	−0.4488
601–800	1.2	0.5	0.3	−0.3927	−1.0472	−0.4488

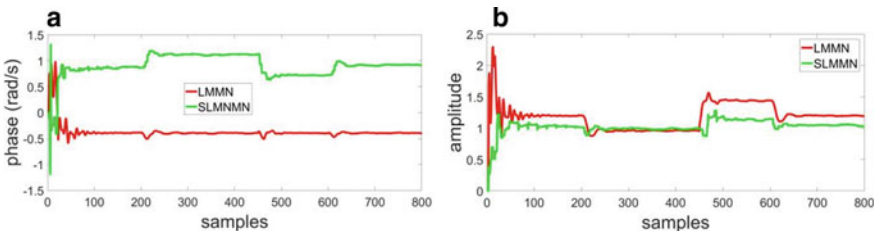


Fig. 2 **a** Comparison of different algorithms based on fundamental phase. **b** Comparison of different algorithms based on fundamental amplitude

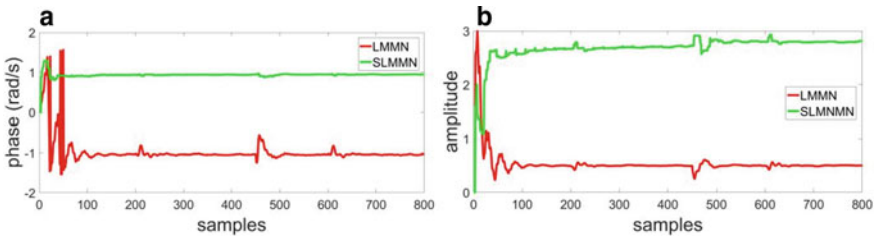


Fig. 3 **a** Comparison between the phase of third harmonic based on different algorithms. **b** Comparison between the amplitude of third harmonic based on different algorithms

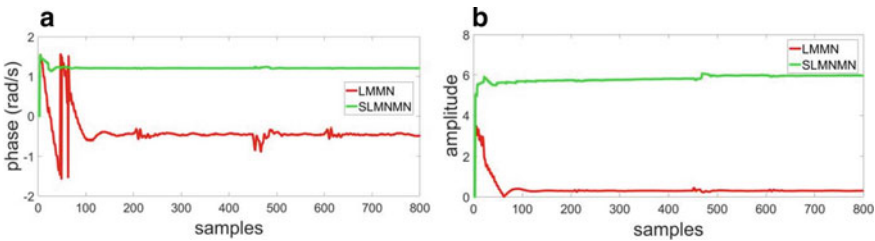


Fig. 4 **a** Comparison between the phase of fifth harmonic based on different algorithms. **b** Comparison between the amplitude of fifth harmonic based on different algorithms

Table 2 Different harmonic analyzer error values under 20, 30, and 40 dB noise [4]

SNR (dB)	LMMN			SLMMN		
	20 dB	30 dB	40 B	20 dB	30 dB	40 B
a1 err	0.5891	0.2082	0.3654	0.2359	0.1550	0.3411
a3 err	0.2138	0.0407	0.3769	0.2062	0.2280	0.4214
a5 err	0.3564	0.2328	0.4615	0.4564	0.3799	0.0356

4 Conclusions

LMMN and SLMMN algorithms enhance the estimation of signals and are more effective than the basic LMS in achieving the desired signal. Hence, resolve the critical issue for the smart grid phasor estimation.

References

1. Y. Xia, S.C. Douglas, D.P. Mandic, Adaptive frequency estimation in smart grid applications: exploiting noncircularity and widely linear adaptive estimators. *IEEE Signal Process. Mag.* **29**(5), 44–54 (2012)

2. M.G. Moreira, A.S. Vale-Cardoso, C.H. Martins, M.C. Leles, Enhanced LMS for online frequency estimation under wide spectrum disturbances, in *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, (2019), pp. 0530–0534
3. S. Nanda, P.K. Dash, A Gauss–Newton ADALINE for dynamic phasor estimation of power signals and its FPGA implementation. *IEEE Trans. Instrum. Meas.* **67**(1), 45–56 (2016)
4. R.N. Shaw, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1) (2019)
5. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON) held at Deenbandhu Chhotu Ram University of Science and Technology, SONEPAT, India on 28 Feb–1 Mar 2020*
6. Y. Belkhier, A. Achour, N. Ullah, R.N. Shaw, Modified passivity-based current controller design of permanent magnet synchronous generator for wind conversion system. *Int. J. Model. Simul.* (2020). <https://doi.org/10.1080/02286203.2020.1858226>
7. P. Rakesh, T.K. Kumar, F. Albu, Modified least-mean mixed-norm algorithms for adaptive sparse system identification under impulsive noise environment, in *2019 42nd International Conference on Telecommunications and Signal Processing (TSP)* (2019), pp. 557–561
8. D.G. Lee, S.H. Kang, S.R. Nam, Modified dynamic phasor estimation algorithm for the transient signals of distributed generators. *IEEE Trans. Smart Grid* **4**(1), 419–424 (2012)

Design and Analysis of Brain-Implantable Antenna for Neural Signal Transmission



Vipan Kumar Gupta and Payal Mehra

Abstract In this paper a biocompatible, wide band antenna is designed for brain-implantable devices. Antenna is characterized to operate in seven-layer complex brain structure. Antenna has center resonance at 2.4 GHz and wide bandwidth of 450 MHz that makes antenna enable to cover medical body area network (MBAN) and industrial scientific and medical (ISM) bands. Antenna size is miniaturized by open-end slotting, and broadside radiation pattern is achieved with the directivity of 4.34 dBi in the off-side direction (away from human head). Antenna has low specific absorption rate (SAR) of 0.044 W/kg. Results show that antenna can be used for brain-implantable devices.

Keywords Brain implantable · ISM · WBAN · Broadside radiation · Seven-layer brain tissue

1 Introduction

Technological advancements in wireless implantable devices are enabling the treatment and controlling of brain disorders caused by spinal cord and traumatic brain injury and strokes. Brain-implantable devices are providing solution for brain interfacing to external machines (like computer, wheelchair) and stimulation to the neurons (for controlling function of urinary bladder system or tongue, etc.) for the patients suffering with Alzheimer's, paralysis, and other related diseases [1–3]. Implantable bio-sensors can extract the intracranial signal like electrocardiography (ECG) and intracranial pressure, and these signals are used to control and monitor the functionality of brain. Establishing a reliable wireless communication link from inside the human brain to the free space environment is one of the major challenges for the deployment of brain-implantable system. Implementation of wireless brain-implanted system is reported in [4–7]. Antenna with wide bandwidth and broadside type radiation characteristics is must to set up the reliable communication link.

V. K. Gupta (✉) · P. Mehra

Electronics and Communication Engineering Department, Sri Sai College of Engineering and Technology, Badhani, Pathankot, India

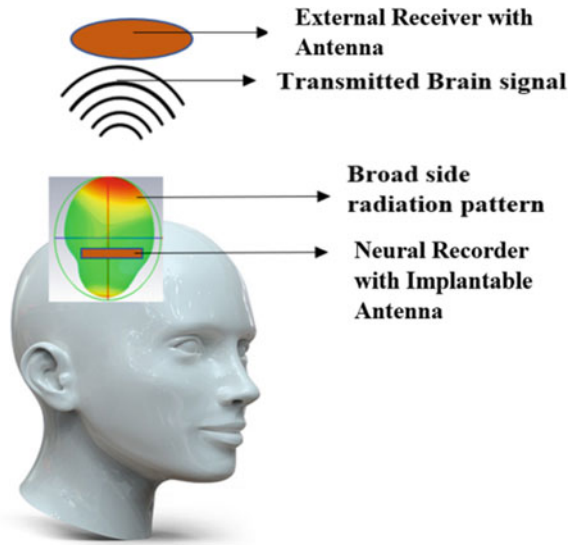


Fig. 1 Schematic view of the implantable antenna's role in wireless brain monitoring system

Schematic view of the wireless brain-monitoring system is shown in Fig. 1. A neural sensor module integrated with implantable antenna is placed inside the head. Neural signals like CSF pressure, temperature, and variation of tissue effective permittivity are transmitted through broadside electromagnetic radiations to the external receiver antenna for data processing.

Human head is a complex structure of multiple layers with different dielectric properties and thickness, and it creates unfavorable conditions for the transmission of electromagnetic waves [6]. Therefore, designing an implantable antenna has its own challenges like frequency detuning, impedance matching, miniaturization, broadside radiation pattern, and ensuring tissue safety from near-field radiations [8]. In [9–11], antenna is miniaturized through stack patch technology and open-end slots. However, omnidirectional radiation characteristics are displayed by these structures which is an undesirable feature for implantable antenna. Back lobe power may cause heating of brain tissue and affect adversely. Double split ring loaded loop antenna is designed in [12] but suffers with extremely low gain. Frequency detuning and impedance mismatch effects are reduced by making antenna biocompatible. In this concern, antenna is shielded with a superstrate layer or coated with low loss materials to evade direct contact of antenna's radiator with brain tissue. Zirconia ($\epsilon_r = 29$, loss $\tan \approx 0$) [13] and Silastik MDX-4210 Elastomer ($\epsilon_r = 3.3$, loss $\tan \approx 0$) [14] are commonly used material for antenna coating. High-permittivity material like RO 3010 ($\epsilon_r = 10.2$) [15] and Taconic RF-35 ($\epsilon_r = 3.5$) [16] are used as superstrate layer. Implantable antennas designed in [17–19] have smaller geometry, good radiation property, and biocompatible. However, these structures are designed for low implant depth of less than 5 mm.

Based on the above issues, present authors have designed a miniaturized antenna that can be implanted at the depth of 12 mm in the middle of Dura and CSF layers of brain tissue. A superstrate layer is used to insulate antenna with conducting body tissue and avoiding impedance distortion. Paper is organized as: Sect. 2 describes the properties of brain tissue and antenna design and behavior. In Sect. 3, antenna performance is evaluated in terms of reflection coefficient, efficiency, power loss in body tissue, and radiation pattern. Finally, in Sect. 4, SAR value is evaluated to ensure the tissue safety from the electromagnetic radiations.

2 Brain Tissue Modeling and Antenna Configuration

It is difficult to characterize the designed antenna with the accurate human brain. Therefore, a simplified seven-layer numerical model is used for the simulation of antenna. Planar area of brain simulation model is 80 mm × 80 mm, with thickness of 72.2 mm. Dimensions of brain model are large enough to avoid diffraction from edges. Material properties and thickness of layers are considered from the data given in [20–23]. Properties of brain tissue layers are given in Table 1. Figure 2 shows the simulation setup in which antenna is implanted between Dura and CSF layers. The cerebral spinal fluid (CSF) is a protective layer that covers the brain cortex and spine and can reflect some of the pathological and physiological changes that occur in the brain. Thus, most of the implantable devices are placed in that region [24].

Geometry of the designed antenna is shown in Fig. 3a, and value of geometrical parameters is given in Table 2. A coplanar waveguide (CPW) feed antenna is designed on thin substrate of RO 3010 with thickness of 0.25 mm, relative permittivity of 10.2, and loss tan of 0.0022. In CPW, technology radiator and ground share common surface. Planar dimensions of antenna are 8.4 mm × 10 mm which are equal to $0.0672\lambda_0 \times 0.08\lambda_0$, where λ_0 is the free space wavelength at 2.4 GHz. Open-ended vertical slots are etched on the radiator to tune the antenna at desired resonance of 2.4 GHz. Etching of slots has increased the electric current length path that helped to reduce the size of antenna. Surface current plot in Fig. 3b represents the current path trace by antenna at 2.4 GHz. It is clear that maximum current is concentrated on

Table 1 Brain tissue properties at 2.45 GHz [20, 21]

Tissue layer	Thickness (mm)	Dielectric constant	Loss tangent	Tissue layer	Thickness (mm)	Dielectric constant	Loss tangent
Skin	1	42.923	0.273	CSF	2	66.319	0.385
Fat	2	5.285	0.145	White Matter	3.7	48.994	0.271
Cortical Bone	7	11.410	0.252	Gray Matter	55	36.226	0.246
Dura	1.5	42.099	0.292	–	–	–	–

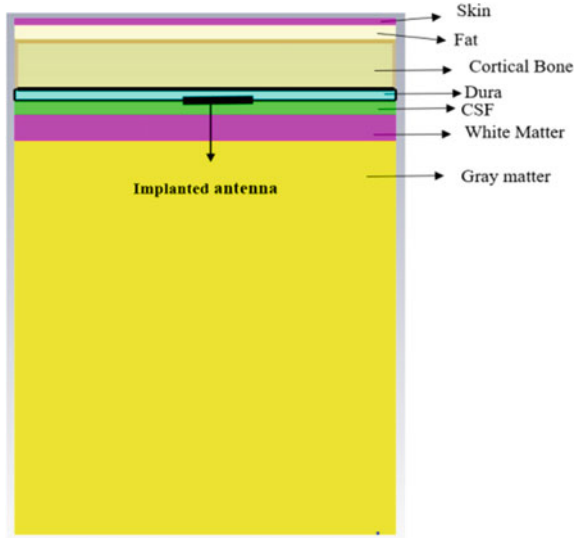


Fig. 2 Simulation setup showing the position of implanted antenna inside the seven-layer brain tissue

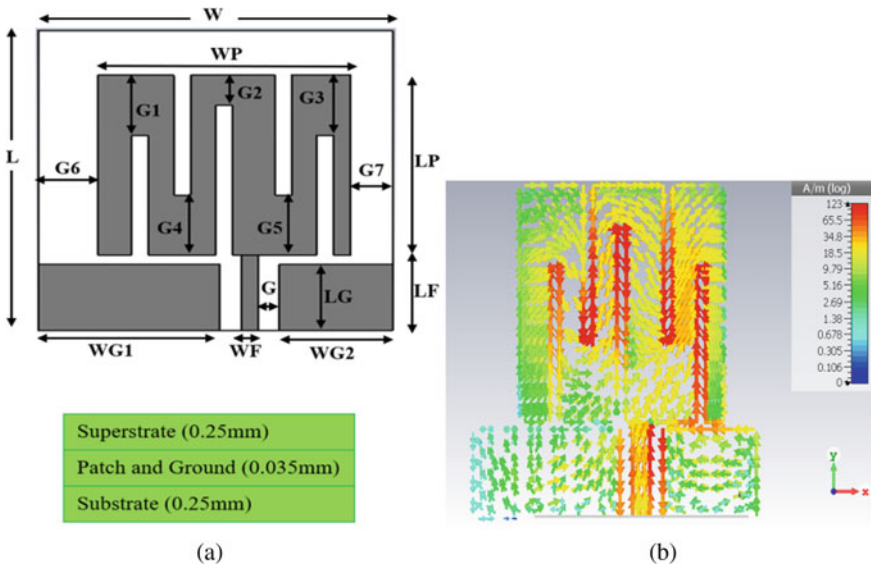


Fig. 3 a Geometry of the designed antenna with cross-sectional view, b surface current distribution at 2.4 GHz

Table 2 Geometrical parameters of the designed antenna

Parameter	Value (mm)	Parameter	Value (mm)	Parameter	Value (mm)	Parameter	Value (mm)
W	8.4	L	10	G	0.4	G4	2
WF	0.4	LF	2.5	G1	2	G5	2
WG1	4.5	LP, WP	6.0	G2	1	G6	1.4
WG2	2.7	LG	2.2	G3	2	G7	1

the edges of slot. Length of the slots is chosen to maintain the impedance matching. In addition to this, offset fed is used to maintain the impedance of feed line. Feed is offset by 1.2 mm to maintain 50-Ω impedance for a thin substrate [25–27].

3 Results and Discussion

Reflection coefficient plot of proposed antenna is shown in Fig. 4. It shows that antenna has covered the wide bandwidth from 2.2165 to 2.6675 GHz with reference to the center frequency of 2.3967 GHz. It has -10 dB impedance bandwidth of 450 MHz. In this, bandwidth antenna has effectively covered the MBAN band ranging from 2.36 to 2.4 GHz and ISM band from 2.4 to 2.48 GHz. Reflection coefficient at center resonance is -26.532 dB which shows that antenna has good impedance matching in the brain tissue.

Plot for radiation efficiency is shown in Fig. 5. Both the total and radiation efficiency are 0.65%. This low efficiency is due to the power absorbed by brain tissue. Simulation-based analysis for power loss in brain issue has been performed. Figure 6

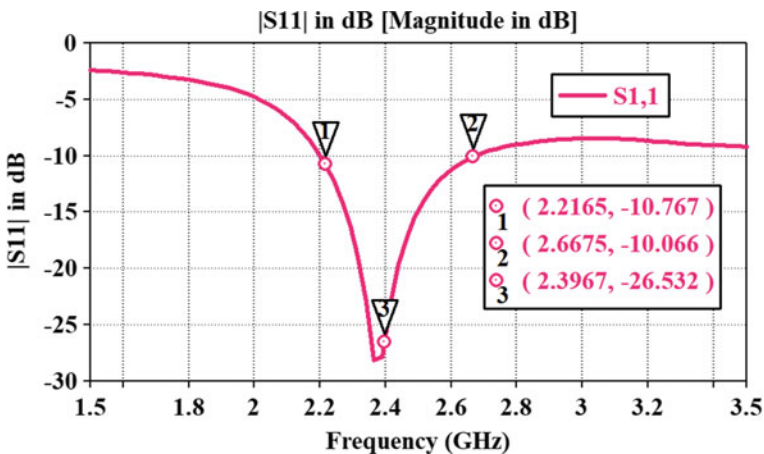


Fig. 4 Reflection coefficient of proposed antenna

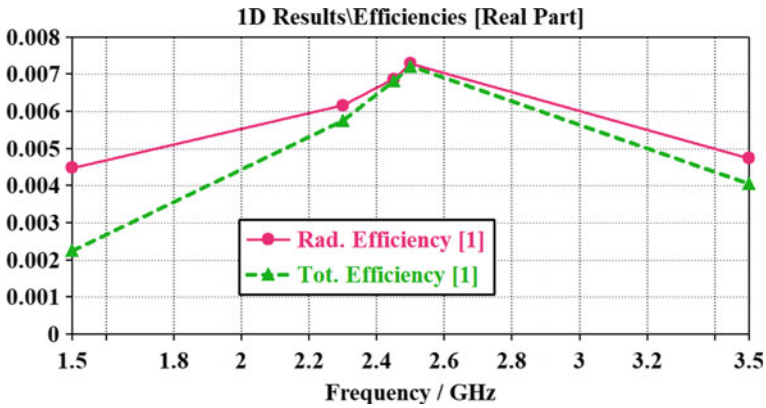


Fig. 5 Plot for efficiency of proposed antenna

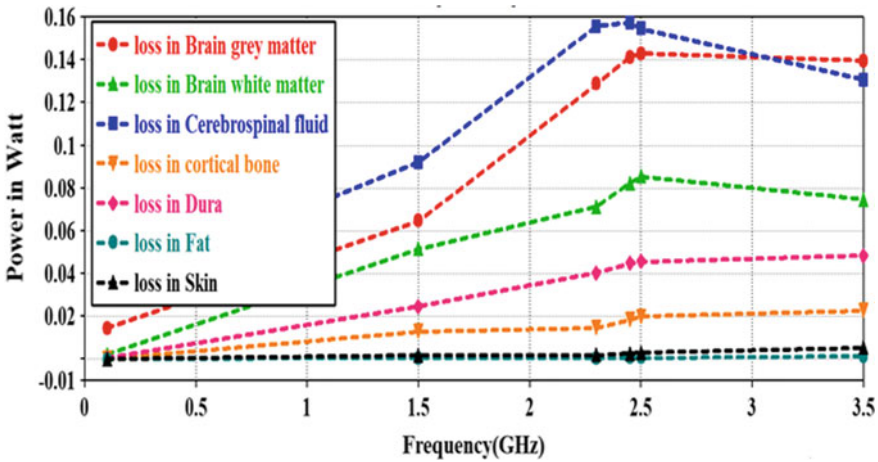


Fig. 6 Plot for power lost in brain tissue layers

shows the plot for power loss in each layer of brain model, and the value of power absorbed at 2.4 GHz is given in Table 3. Antenna is simulated for 0.5 W of input power. Total power loss in tissue and substrate material is 0.451 W. It shows that very low amount of power is lost by antenna structure, impedance mismatch, and surface waves.

3D radiation pattern of antenna is shown in Fig. 7. In Fig. 7b, brain tissue layer is hidden to view the antenna. Both the plots are simulated under same setup. High directivity of 4.34 dBi is achieved at 2.4 GHz. 2D plot in Fig. 8 shows that broadside radiation pattern is achieved for both the E-plane and H-plane. Broadside radiation characteristics is the most desirable feature of the implantable antenna for transmitting maximum radiated power to the external monitoring device.

Table 3 Value of power lost at 2.4 GHz

Layers	Power loss (W)
Skin	0.002
Fat	0.0005
Cortical Bone	0.01432
Dura	0.0358197
CSF	0.1394
White Matter	0.064930
Gray Matter	0.11108
RO 3010	0.0043
Total loss	0.4510

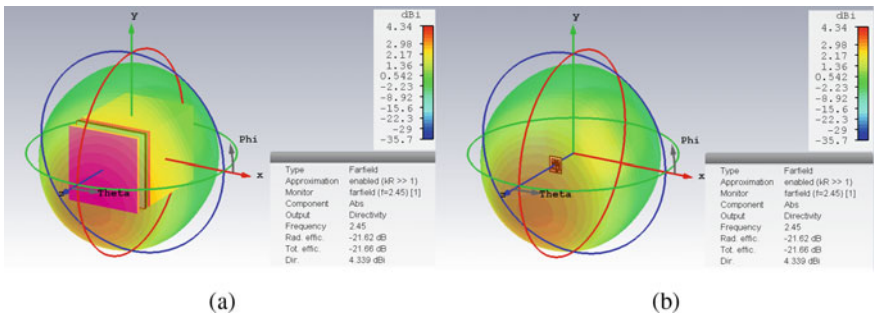


Fig. 7 Directivity plot **a** shown with brain tissue, **b** brain tissue hidden

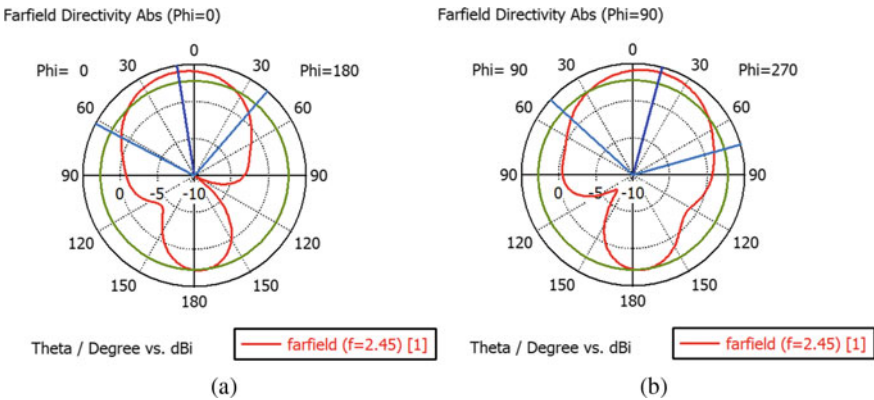


Fig. 8 Radiation plot **a** E-plane, **b** H-plane

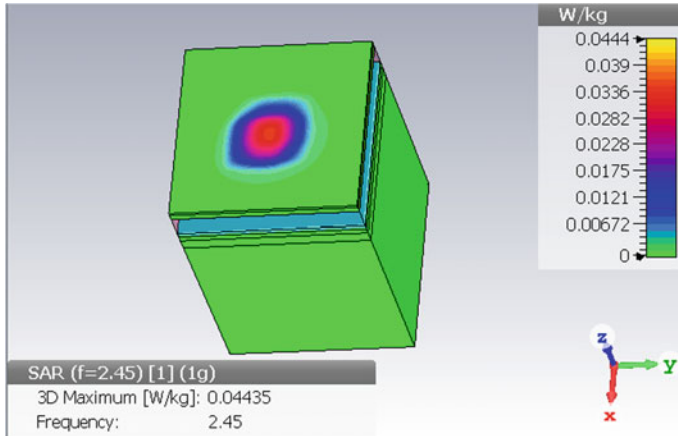


Fig. 9 Specific absorption rate at 2.45 GHz

4 Specific Absorption Rate Value

Specific absorption rate is the safety measure of the body tissue from electromagnetic radiations of implanted antenna. SAR value should be less than 2 W/kg for 10 g of body tissue mass [25]. Plot for SAR value of proposed antenna is shown in Fig. 9 for the proposed structure SAR input power of 25 μ W (IEEE safety regulation).

5 Conclusions

A CPW fed biocompatible implantable antenna is designed and characterized to implant in the human brain at the depth of 12 mm. Antenna is a compact low profile structure of 8.4 mm \times 10 mm \times 0.5 mm. Electric current length is enhanced by etching multiple slots and impedance matching is achieved with offset feed. Analysis for low efficiency is done through the power lost data in brain tissue layers. Antenna has wide band performance, broadside radiation pattern, and low specific absorption rate that makes antenna suitable for transmitting signal from the brain-implanted devices.

References

1. C.W.L. Lee, A. Kiourt, J. Chae, J.L. Volakis, A high-sensitivity fully passive neurosensing system for wireless brain signal monitoring. *IEEE Trans. Microw. Theory Technol.* **63**(6), 2060–2068 (2015)

2. G. Schalk, E.C. Leuthardt, Brain-computer interfaces using electrocorticographic signals. *IEEE Rev. Biomed. Eng.* **4**, 140–154 (2011)
3. A.T. Connolly et al., A novel lead design for modulation and sensing of deep brain structures. *IEEE Trans. Biomed. Eng.* **63**(1), 148–157 (2016)
4. T. Seo, S. Oh, D. Jung, Y. Huh, J. Cho, Y. Kwon, Noninvasive brain stimulation using a modulated microwave signal. *J. Electromagn. Eng. Sci.* **18**(1), 70–72 (2018)
5. W.-C. Chen, C.W. L. Lee, A. Kiourti, J.L. Volakis, A multi-channel passive brain implant for wireless neuropotential monitoring. *IEEE J. Electromagn., RF, Microw. Med. Biol.* **2**(4), 262–269 (2018)
6. L. Song, Y. Rahmat-Samii, An end-to-end implanted brain-machine interface antenna system performance characterizations and development. *IEEE Trans. Antennas Propag.* **65**(7), 3399–3408 (2017)
7. A. Kiourti, C.W.L. Lee, J. Chae, J.L. Volakis, A wireless fully passive neural recording device for unobtrusive neuropotential monitoring. *IEEE Trans. Biomed. Eng.* **63**(1), 131–137 (2016)
8. A. Kiourti, K.S. Nikita, A review of implantable patch antennas for biomedical telemetry: challenges and solutions [wireless corner]. *IEEE Antennas Propag. Mag.* **54**(3), 210–228 (2012)
9. A. Kiourti, J.R. Costa, C.A. Fernandes, A.G. Santiago, K.S. Nikita, Miniature implantable antennas for biomedical telemetry: From simulation to realization. *IEEE Trans. Biomed. Eng.* **59**(11), 3140–3147 (2012)
10. M. Lee, T.C. Yo, C.H. Luo, C.H. Tu, Y.Z. Juang, Compact broadband stacked implantable antenna for biotelemetry with medical devices. *Electron. Lett.* **43**(12), 660–662 (2007)
11. W.-C. Liu, F.-M. Yeh, M. Ghavami, Miniaturized implantable broad-band antenna for biotelemetry communication. *Microw. Opt. Technol. Lett.* **50**(9), 2407–2409 (2008)
12. S. Ma, T. Björninen, L. Sydänheimo, M.H. Voutilainen, L. Ukkonen, Double split rings as extremely small and tuneable antennas for brain implantable wireless medical microsystems. *IEEE Trans. Antennas Propag.* (2020)
13. A.K. Skrivervik, F. Merli, Design strategies for implantable antennas, in *Proceedings of the Antennas and Propagation Conference, Loughborough, UK*, 2011
14. T. Karacolak, R. Cooper, J. Butler, S. Fisher, E. Topsakal, In Vivo verification of implantable antennas using rats as model animals. *IEEE Antennas Wirel. Propag. Lett.* **9**, 334–337 (2010)
15. S. Hout, J.-Y. Chung, Design and characterization of a miniaturized implantable antenna in a seven-layer brain phantom. *IEEE Access* **7**, 162062–162069 (2019)
16. V.K. Gupta, D. Thakur, Design and performance analysis of a CPW-fed circularly polarized implantable antenna for 2.45 GHz ISM band. *Microw. Opt. Technol. Lett.* (2020) (accepted)
17. B. Rana et al., An implantable antenna with broadside radiation for a brain machine interface. *IEEE Sensors J.* **19**(20), 9200–9205 (2019)
18. S.A.A. Shah, H. Yoo, Scalp-implantable antenna systems for intracranial pressure monitoring. *IEEE Trans. Antennas Propag.* **66**(4), 2170–2173 (2018)
19. M. Saxena, R.N. Shaw, J.K. Verma, A novel hash-based mutual RFID tag authentication protocol. In *Advances in Intelligent Systems and Computing*, vol. 847 (Springer, Berlin, 2019), pp. 1–12. https://doi.org/10.1007/978-981-13-2254-9_1
20. S. Gabriel, R. Lau, C. Gabriel, The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz. *Phys. Med.* **41**(11), 2251–2269 (1996). Available <https://niremf.ifac.cnr.it/tissprop/>
21. I. Das, R.N. Shaw, S. Das, Location-based and multipath routing performance analysis for energy consumption in wireless sensor networks, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh, vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_59
22. X.Y. Liu, Z.T. Wu, Y. Fan, E.M. Tentzeris, A miniaturized CSRR loaded wide-beamwidth circularly polarized implantable antenna for subcutaneous real-time glucose monitoring. *IEEE Antennas Wirel. Propag. Lett.* **16**, 577–580 (2017)
23. A. Drossos, V. Santomaa, N. Kuster, The dependence of electro-magnetic energy absorption upon human head tissue composition in the frequency range of 300–3000 MHz. *IEEE Trans. Microw. Theory Technol.* **48**(11), 1988–1995 (2000)

24. H. Bahrami, B. Gosselin, L.A. Rusch, Realistic modeling of the biological channel for the design of implantable wireless UWB communication systems, in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug./Sept. 2012, pp. 6015–6018
25. S. Gabriel, R.W. Lau, C. Gabriel, The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues. *Phys. Med. Biol.* **41**(11), 2271–2293 (1996)
26. A. Lleo et al., Cerebrospinal fluid biomarkers in trials for Alzheimer and Parkinson diseases. *Nat. Rev. Neurol.* **11**(1), 41–55 (2015)
27. IEEE Standard for Safety Levels with Respect to Human Exposure to Radiofrequency Electromagnetic Fields, 3 kHz to 300 GHz, IEEE Standard C95.1, 2005

A Novel Design of FSS-Based Absorber Integrated Microstrip Antenna



Priyanka Das and Kaushik Mandal

Abstract In this article, the design of a dual-band FSS-based absorber with high angular stability has been proposed to function from 3.4–4.6 GHz to 7.9–9 GHz. An air gap of 1.6 mm is kept between two identical layers of concentric split ring resonator based FSS. The bottom FSS has a metallic backing to ensure minimum transmission. A square patch-based microstrip antenna has been designed to operate at 6.1 GHz. When this antenna is integrated with the absorber, out of band RCS reduction of the antenna takes place for a wide range of frequencies.

Keywords FSS · Absorber · Microstrip antenna

1 Introduction

Microwave absorbers are used to absorb the impending electromagnetic waves by inhibiting their transmission and reflection through the surface. They are used in stealth technology for radar cross-sectional reduction by minimizing the scattered power. Lumped resistors, resistive ink printed structures, and metamaterials are used to design absorbers. Since metamaterial absorbers have subwavelength dimensions, they facilitate realization of low-profile easy-to-fabricate compact structures. In the recent years, a lot of research has been conducted for realization of metamaterial based absorbers. For terahertz applications, a single-layer absorber has been proposed in [1] which functions at two different frequency bands. In [2], a multilayered metamaterial structure has been proposed which comprises six concentric closed rings. An omnidirectional electromagnetic absorber [3] composed of a gradient index metamaterial shell which has non-resonant I-shaped structures and a lossy dielectric core has been reported. In [4], a metamaterial absorber based on sun fractal structure operating at two distinct frequencies has been investigated. A reconfigurable absorber [5]

P. Das (✉)
University of Engineering and Management, Kolkata, India

K. Mandal
Institute of Radio-Physics, University of Calcutta, Kolkata, India

has been implemented by employing varactors. 3D stacked metamaterial array of split ring resonators (SRR) has been utilized for rectification of microwave energy as demonstrated in [6]. In this paper, a compact dual-band FSS-based absorber is designed without using lumped resistors or resistive sheet.

2 Design of Absorber

The unit cell of the absorber ($0.1\lambda \times 0.1\lambda$) comprises two layers of three metallic concentric square loops split at the center of the side arms on FR4 substrate having thickness of 1.6 mm separated by an air gap of 1.6 mm. λ is the wavelength corresponding to the lowest resonant frequency of the absorber. The top view and bottom view of each unit cell of the absorber are illustrated in Fig. 1. The reflection coefficient of the absorber varies with frequency as shown in Fig. 2. Dual-band absorption takes places at frequencies (3.4–4.6 GHz) and (7.9–9 GHz) at which reflection coefficient is minimum since transmission coefficient is zero due to metallic backing as shown in Fig. 3.

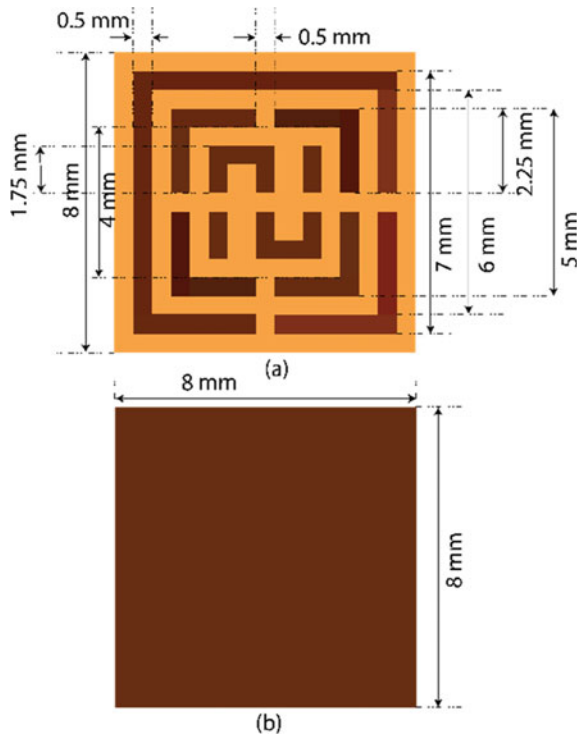


Fig. 1 Unit cell of the absorber **a** top view (first layer and second layer), **b** bottom view

Fig. 2 Reflection coefficient characteristics

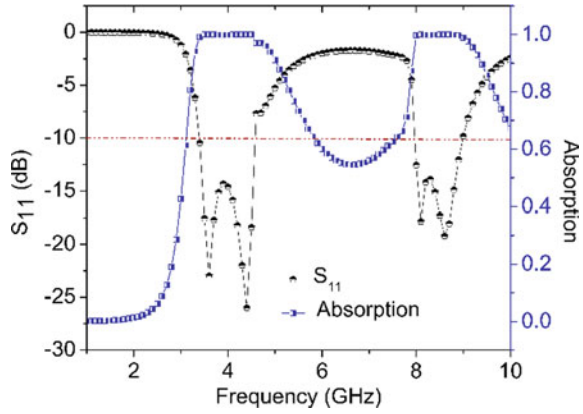
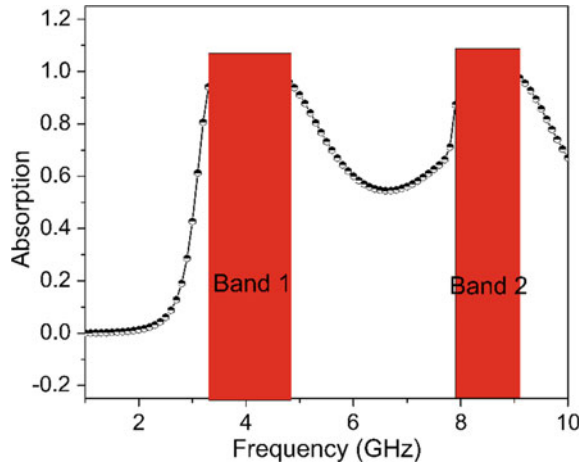


Fig. 3 Absorption characteristics of the unit cell



$$\text{Absorption} = 1 - |S_{11}|^2 - |S_{21}|^2 \tag{1}$$

$$\text{When } S_{21} = 0, \text{ Absorption} = 1 - |S_{11}|^2 \tag{2}$$

For maximizing absorption, the magnitude of S_{11} must be minimized. The effective impedance of the absorber is obtained from the following equation

$$Z_{\text{eff}} = \frac{1 + S_{11}}{1 - S_{11}} \tag{3}$$

The absorber exhibits magnetic resonance at 3.6 and 4.4 GHz with anti-directional surface currents on the top FSS layer and the ground layer. On the other hand, electric resonance is registered at 8.2 and 8.6 GHz with surface currents flowing in the same

direction on the top FSS layer and the ground layer. Inductive coupling takes place between the top FSS layer and the bottom FSS layer at 3.6, 8.2, and 8.6 GHz with currents flowing in the same direction while capacitive coupling takes place between the top and bottom FSS layers at 4.4 GHz with currents flowing in the opposite direction.

It is shown in Fig. 4 that the effective impedance is 1 at the absorption bands, since reflection coefficient is nearly zero. This happens when the input impedance of the absorber matches with the intrinsic impedance of air. In absence of the air gap, a variation in the reflection coefficient characteristics if the absorber is observed, as evident from Fig. 5. The absorption bandwidth is narrowed when the air gap between

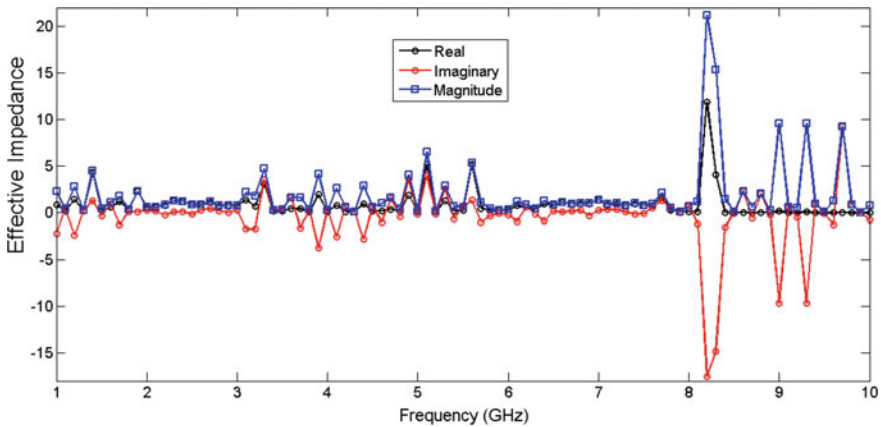


Fig. 4 Plot of effective impedance versus frequency

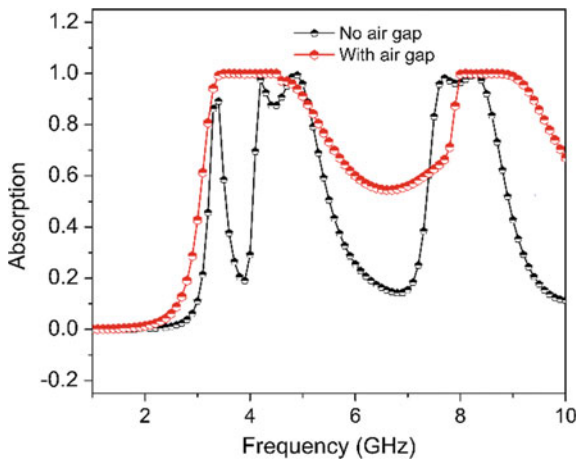


Fig. 5 Comparison plot of absorption characteristics of the unit cell with and without air gap

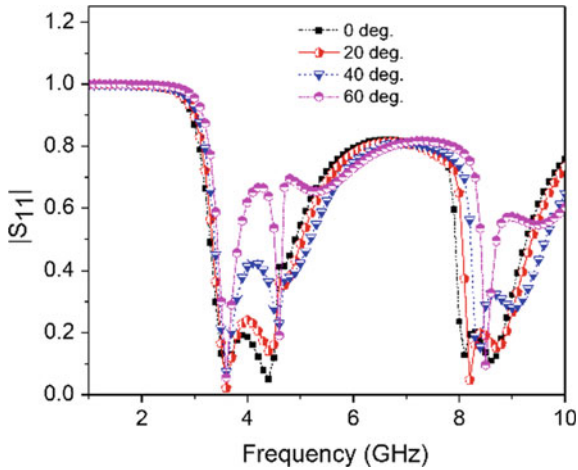


Fig. 6 Angular stability in reflection coefficient characteristics of the absorber

the FSS layers is absent. This illustrates that the introduction of a small air gap in the absorber helps to acquire matching of the surface impedance of the absorber with that of air over a wide band. The reflection coefficient characteristics of the unit cell of the absorber exhibits high angular stability till 60° as depicted in Fig. 6 due to its symmetrical structure.

3 Design of Microstrip Patch Antenna

A square patch ($0.2\lambda \times 0.2\lambda$) is designed on FR4 substrate having a thickness of 3.2 mm (0.065λ) with full metallic ground at the bottom. λ corresponds to the resonant frequency (6.1 GHz) of the antenna. The coaxial feed point is 2 mm from the center of the patch as shown in Fig. 7.

4 Integration of Absorber with Antenna

Structural mode RCS reduction of the antenna can be accomplished by integrating it with the microwave absorber. The integrated structure is designed in such a way that the absorber encompasses the square patch at the center from all four sides. The top FSS layer of the absorber is placed at a vertical distance of 1.6 mm from the antenna substrate. The bottom FSS layer is merged with the antenna substrate such that the absorber metallic patch and the microstrip patch lie on the same plane as shown in Fig. 8. Out of band monostatic RCS reduction of the antenna takes place over a wide range of frequencies from 2.5–5.0 GHz to 5.5–10 GHz. Maximum monostatic

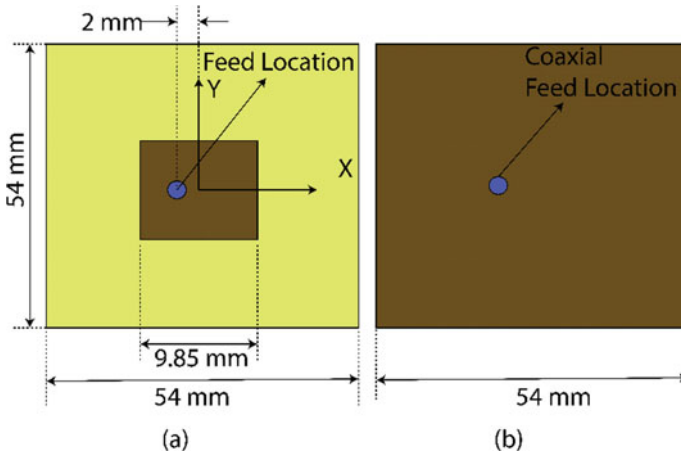


Fig. 7 Structure of the microstrip patch antenna a top view, b bottom view

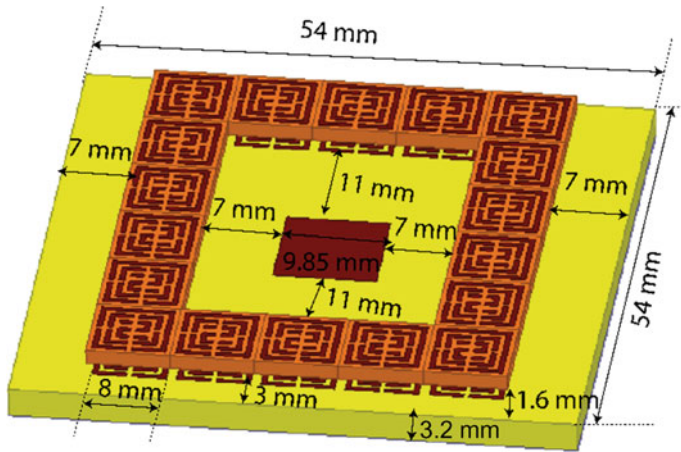


Fig. 8 Integration of the antenna with the absorber

RCS reduction of the antenna on integration with absorber is 11.5 dB at 3.9 GHz as shown in Fig. 9. Bistatic RCS reduction of the antenna on integration with absorber is 11.5 dB at 0° at 3.9 GHz as shown in Fig. 10.

5 Conclusion

In this article, wideband RCS reduction of coaxial-fed microstrip patch antenna has been demonstrated by integrating it with a dual-band absorber in order to create a

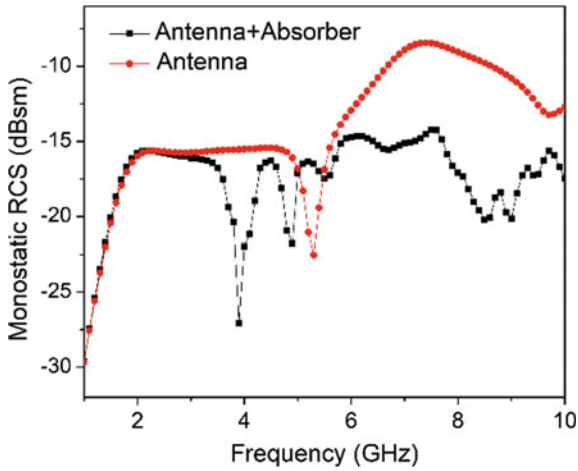


Fig. 9 Monostatic RCS reduction of the antenna with the absorber

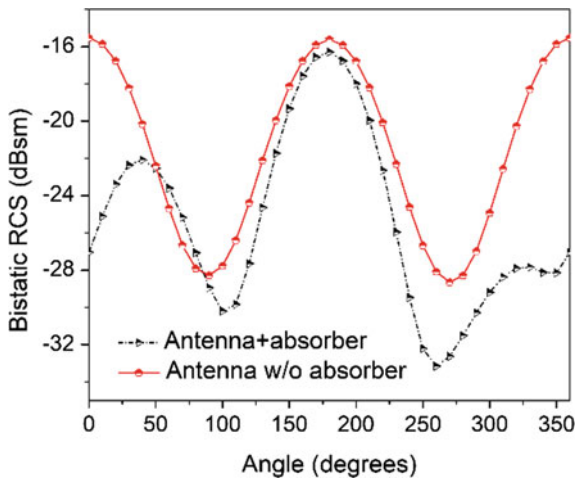


Fig. 10 Bistatic RCS reduction of the antenna with the absorber at 3.9 GHz

composite structure. The absorber has been designed using two identical layers of closely spaced metallic square loops split at the center of their sides in order to match the surface impedance with that of intrinsic impedance of air by creating resonances at 3.6 GHz, 4.4 GHz, 8.2 GHz, and 8.6 GHz, respectively. The proposed antenna is a suitable candidate for stealth applications.

References

1. J. Wu, W. Qin, Terahertz dual-band nearly perfect absorbers based on combined of two types of FSS elements, in *2012 IEEE MTT-S International Microwave Workshop Series on Millimeter Wave Wireless Technology and Applications*, Nanjing, 2012, pp. 1–3
2. L. Huang, H. Chen, Multi-band and polarization insensitive metamaterial absorber. *Prog. Electromagn. Res.* **113**, 103–110 (2011)
3. Q. Cheng, T.J. Cui, W.X. Jiang, B.G. Cai, An omnidirectional electromagnetic absorber made of metamaterials. *New J. Phys.* **12**, 063006 (2010)
4. M. Edries, H.A. Mohamed, S.S. Hekal, M.A. El-morsy H.A. Mansour, Simulation design of dual band metamaterial absorber based on the fractal structure, in *2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, Atlanta, GA, USA, pp. 1335–1336, 2019
5. M.M. Bait-Suwailam, T.S. Almoneef, A. Alomainy, A wearable reconfigurable electromagnetic metamaterial absorber using artificial magnetic inclusions, in *2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, Atlanta, GA, USA, 2019, pp. 1623–1624
6. T.S. Almoneef, O.M. Ramahi, A 3-dimensional stacked metamaterial arrays for electromagnetic energy harvesting. *Prog. Electromagn. Res.* **146**, 109–115 (2014)

Off-State Leakage Concern in Scaling Nanowire FETs



Rajiv Ranjan Thakur, Nitin Chaturvedi, and Nidhi Chaturvedi

Abstract In this work, the simulation-based performance comparisons of the Si nanowire FET have been done for gate length scaling from 90 to 32 nm technology node. The study involves the design and optimization of the critical parameters for improved electrostatic control on the channel. The impact of gate length scaling on the off-state leakage current and threshold voltage roll-off concepts has been discussed. The study reports a drain current enhancement of 48.72 and 72.12% for gate length scaling from 90 to 45 nm and 90 nm to 32 nm technology node, respectively. The maximum mobility of the carrier up to $1173.86 \text{ cm}^{-2}/\text{V}$. and $I_{\text{on}}/I_{\text{off}}$ ratio of $\sim 10^9$ has been reported.

Keywords Nanowire · FETs · Electrostatic potential · Quantum effects

1 Introduction

The quest for the improved electrostatic control has been the driving force for the transition from the planar to non-planar FETs [1]. The device scaling follows a trend setup by the IRDS (former ITRS) and suggests that scaling will lead to short channel effects such as SS and DIBL [2, 3]. The emerging devices should address the SCEs and should be compatible with the planar CMOS platform for the successful realization of the device [4]. Over the time new device geometry has been proposed such as DG MOSFET, FinFET, Q-FinFET, nanowire FETs, nanotubes FETs, nanosheets FETs [5–8]. One-dimensional nanostructure such as nanowire is one of the most emerging device structures for the logic-based circuits and systems due to its excellent electrostatic control mechanism which opens up a way for the future nanoelectronics and optoelectronics devices [9]. The major structure advantage of the gate

R. R. Thakur (✉) · N. Chaturvedi

CSIR-Central Electronics Engineering Research Institute, Pilani, India

Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India

N. Chaturvedi

Birla Institute of Technology and Science, Pilani, India

all around nanowire FETs is the optimal electrostatic control on the channel due to the summation of the individual gate capacitance. One of the major bottlenecks while scaling down the gate length is the off-state leakage control [10]. The scaling will lead low-power, low-cost, and high-speed circuits and systems with increased functionality and device density [11, 12]. The major electronics applications involve digital logics, switches, microprocessors, and IoT [13]. One of the major bottlenecks while scaling down the gate length is the off-state leakage control which will lead to the deterioration of the threshold voltage of the device and is a key parameter from circuit designers point of view [14]. Another way to improve the gate control on the channel is to reduce the oxide thickness which will increase the gate capacitance but ultra-thin gate oxide will lead to reliability and increased power consumption [15]. Thus, the device design plays a critical role for the overall improved performance and proper optimization of the parameters is the key to achieve a device with high on-current (I_{on}) and lower off-state leakage current (I_{off}).

In this work, a simulation-based design and optimization of the performance parameter has been carried out for the gate all around Si nanowire FETs for downscaling of gate length from 90 to 32 nm.

2 Design and Simulation Platform

The simulation works have been carried out on Sentaurus TCAD, a commercially available semiconductor device simulator. The 2D cross section of the device has been shown in Figs. 1 and 2. The design was optimized as per the parameters listed in Table 1. The simulation works were optimized for the three different gate lengths of 32 nm, 45 nm, and 90 nm, respectively. The device dimensions were optimized such that there exists a proper mitigation of the SCEs in the device for the overall performance enhancement. The quantization of the carriers near the channel has been studied using the density gradient quantization model.

3 Results and Discussions

The output characteristics of the Si nanowire FET for different gate length 32 nm, 45 nm, and 90 m are shown in Figs. 3, 4, and 5, respectively, for the applied gate bias of 0.5 V, 1.5 V, 2.0 V, and 2.5 V. For the applied low gate bias of 0.5 V, the drain current was found to be 1.55 μA , 20.8 μA , and 71.4 μA for nanowire FET with gate length of 90 nm, 45 nm, and 32 nm, respectively, as shown in Figs. 3, 4, and 5. When the gate bias was increased to 2.5 V, the drain current was 33.15% and 49.63%, respectively, for the gate length scaling from 90 to 45 nm and 90 nm to 32 nm. The maximum drain current was found to be 0.534 mA for 90 nm gate length, 0.711 mA for 45 nm gate length, and 0.799 mA for 32 nm gate length, respectively.

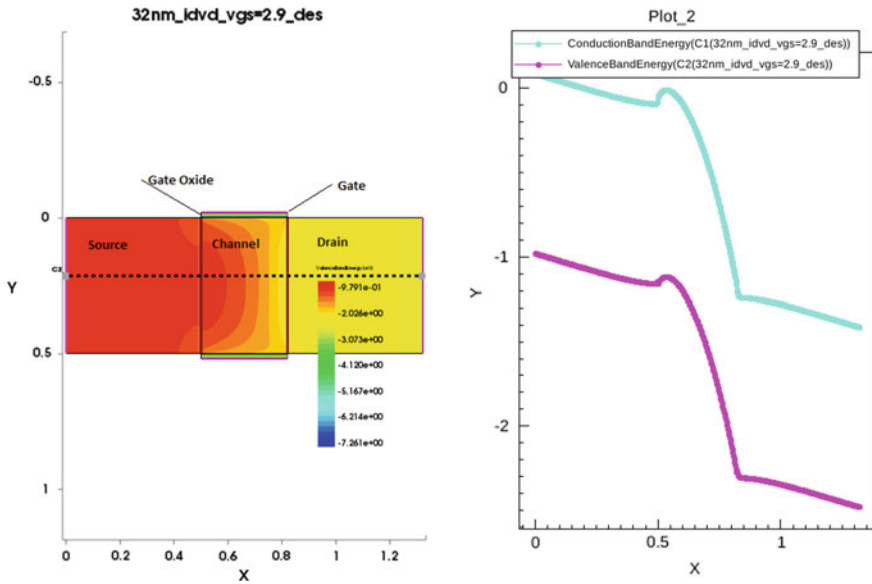


Fig. 1 Simulated device structure 2D cross section on Sentaurus TCAD and its corresponding band diagram

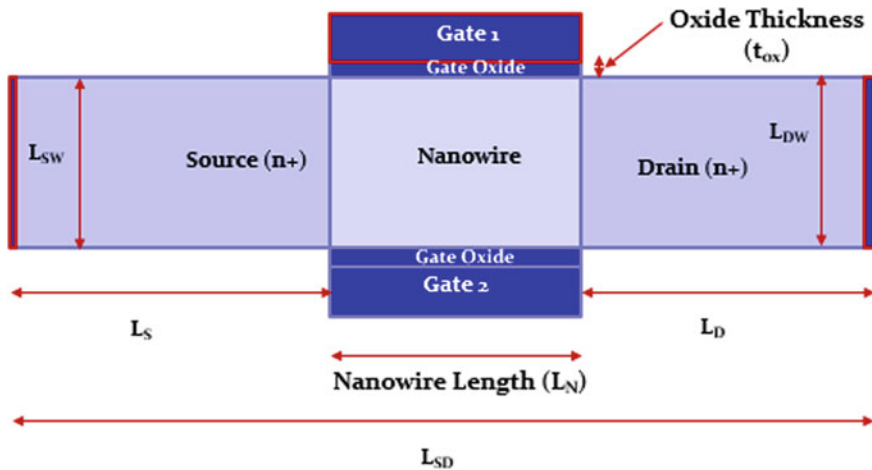


Fig. 2 Lateral 2D drawn cross section of the simulated Si nanowire FET

The transfer characteristics of the simulated Si nanowire FET have been plotted in Fig. 6, and its corresponding logarithmic plot has been depicted in Fig. 7. The on-current (I_{on}) of the device was found to increase as the gate length was scaled from 90 to 45 nm and 90 nm to 32 nm. The maximum drain current for 90 nm Si nanowire FET was found to be 31.2 μA , for 45 nm Si nanowire FET is was 46.4

Table 1 Design and simulation parameters

S. no.	Device parameters	Values
1	Nanowire length (L_N)	32, 45, 90 nm
2	Source/drain length (L_S or L_D)	50 nm
3	Source/drain width (L_{SW} or L_{DW})	30 nm
4	Oxide thickness (t_{ox})	2 nm
5	Gate voltage (V_{gs})	2 V
6	Acceptor doping concentration (N_a)	$1e+16 \text{ cm}^{-3}$
7	Donor doping concentration (N_d)	$1e+18 \text{ cm}^{-3}$
8	Work function (ϕ_{ms})	5.01 eV

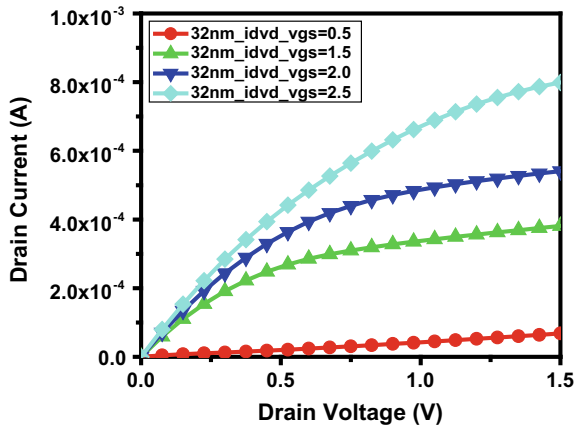


Fig. 3 Output characteristics of Si FET nanowire for 32 nm gate length

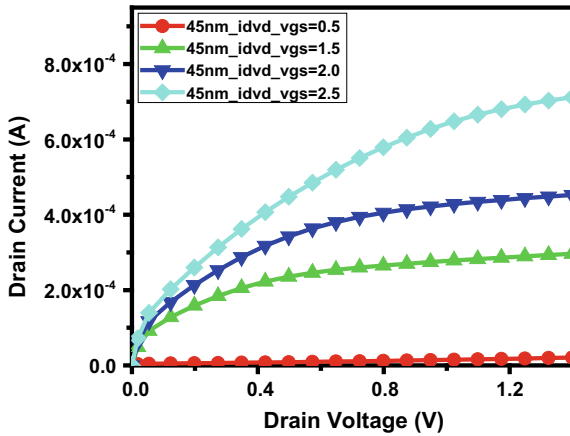


Fig. 4 Output characteristics of Si FET nanowire for 45 nm gate length

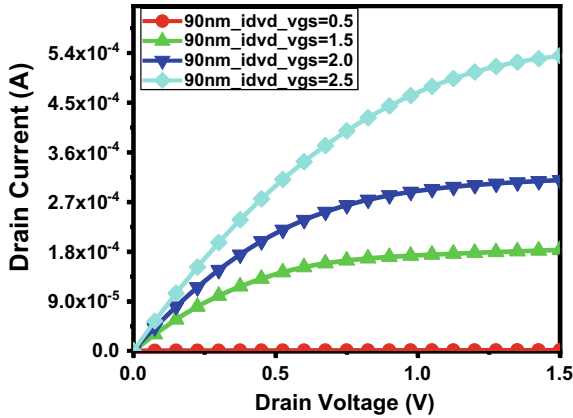


Fig. 5 Output characteristics of Si FET nanowire for 90 nm gate length

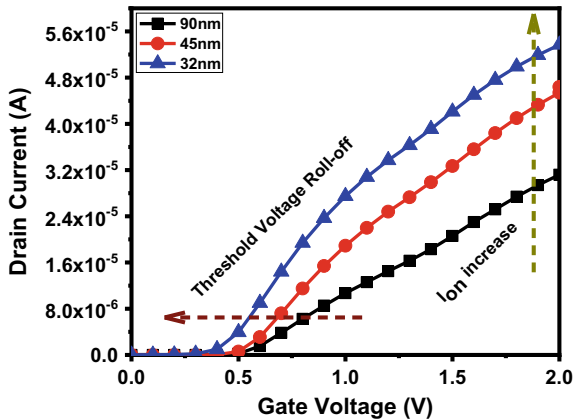


Fig. 6 Transfer characteristics of Si FET nanowire

μA , and for 32 nm Si nanowire FET it was $53.7 \mu\text{A}$. The results show an increment of 48.72% in drain current for gate length scaling from 90 to 45 nm and 72.12% for gate length scaling from 90 to 32 nm, respectively. There was a reduction in the threshold voltage of the device due to the threshold voltage roll-off and the threshold voltage for 90 nm gate length was $\sim 0.6 \text{ V}$, for 45 nm gate length it was $\sim 0.5 \text{ V}$, and for 32 nm gate length it was $\sim 0.4 \text{ V}$. It was also found that as we are scaling down the gate length, one of the major concerns is the control of off-state leakage current in the device. For the simulated device, the off-state leakage current was found to be 26.4 fA for the 90 nm gate length, 27.5 pA for the 45 nm gate length, and 23.9 nA for the 32 nm gate length Si nanowire FET. These values of off-state leakage clearly indicate an increase in off-state leakage as we scaled down the gate length of the device. The scaling of the device is responsible for the short channel effects in the

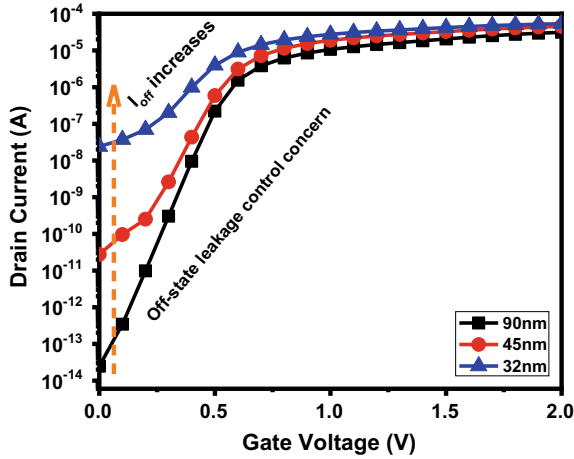


Fig. 7 Logarithmic plot of transfer characteristics of Si FET nanowire

nanoscale devices. The I_{on}/I_{off} ratio of the devices with gate length 90 nm was found to be $\sim 10^9$, $\sim 10^6$ for 45 nm, and $\sim 10^3$ for 32 nm technology node. The mobility of the carriers in the channel was found to be not much reduced as the transport took place in the nanowire FET with gate length of 90 nm as shown in Fig. 8 but for 45 and 32 nm there was a reduction in the mobility due to the scattering effects. It was also observed that less potentials are needed to move a unit charge from one point to another as we move to lower technology transistors as shown in Fig. 9.

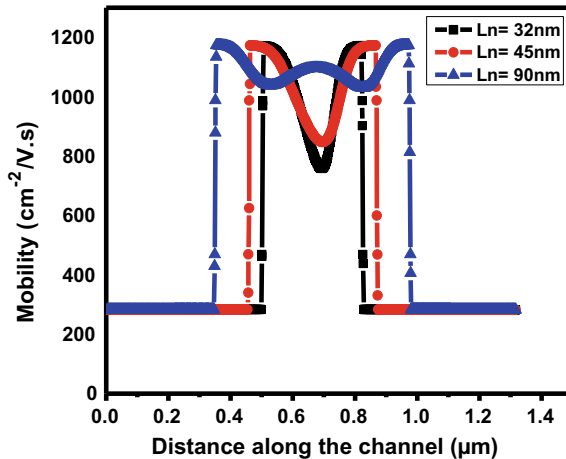


Fig. 8 Mobility comparisons in Si nanowire FET

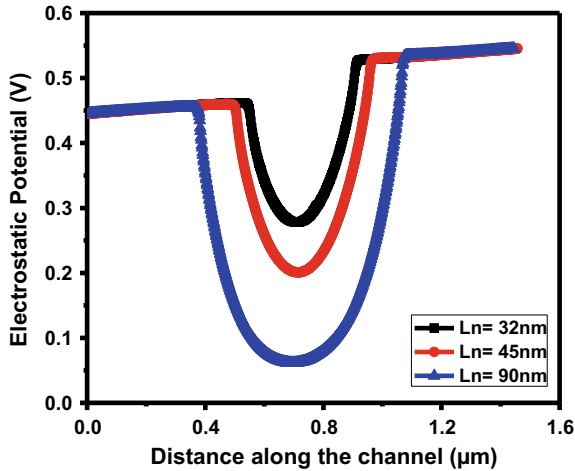


Fig. 9 Electrostatic potential in Si nanowire FET

4 Conclusion

The paper reports off-state leakage concerns in scaling the nanowire FETs and its impact on the performance of the device. It was found that drain current can be enhanced up to 48.72 and 72.12% for gate length scaling from 90 to 45 nm and 90 nm to 32 nm technology node, respectively. The maximum mobility was found to be $1173.86 \text{ cm}^{-2}/\text{V s}$ and $I_{\text{on}}/I_{\text{off}}$ ratio of $\sim 10^9$. The on-current (I_{on}) increases as we scaled down the device but it also increases the off-state leakage (I_{off}) so a trade-off should be made for overall optimized device.

References

1. I. Ferain, C.A. Colinge, J.P. Colinge, Multigate transistors as the future of classical metal–oxide–semiconductor field-effect transistors. *Nature* **479**(7373), 310–316 (2011). <https://doi.org/10.1038/nature10676>
2. B. Hoefflinger, ITRS: the international technology roadmap for semiconductors. Chips 2020 (Springer, Berlin, 2011), pp. 161–174
3. M. Moore, International roadmap for devices and systems (2020)
4. A. Chen, J. Hutchby, V. Zhimov, G. Bourianoff (eds.), *Emerging Nanoelectronic Devices* (Wiley, New York, 2014)
5. R.R. Thakur, P. Singh, Study of carrier scattering and quantization effects in steep retrograded double gate FinFETs for nano technology applications. *Mater. Today: Proc.* **24**, 2019–2023 (2020)
6. R.R. Thakur, P. Singh, Q-FinFET: the next generation Fin FET. *J. Nanoelectron. Optoelectron.* **14**(1), 92–98 (2019). <https://doi.org/10.1166/jno.2019.2444>
7. A. Javey, J. Guo, Q. Wang, M. Lundstrom, H. Dai, Ballistic carbon nanotube field-effect transistors. *Nature* **424**(6949), 654–657 (2003)

8. B.J. Obradovic, J.A. Kittl, M.S. Rodder, Crystalline multiple-nanosheet III-V channel FETs. U.S. Patent 9,484,423, issued November 1, 2016
9. J.-P. Colinge, J.C. Greer, J. Greer, *Nanowire Transistors: Physics of Devices and Materials in One Dimension* (Cambridge University Press, 2016)
10. L. Chang, S. Tang, T.-J. King, J. Bokor, C. Hu, Gate length scaling and threshold voltage control of double-gate MOSFETs, in *International Electron Devices Meeting 2000*. Technical Digest. IEDM (Cat. No. 00CH37138). IEEE, 2000, pp. 719–722
11. T. Poiroux, M. Vinet, O. Faynot, J. Widiez, J. Lolivier, T. Ernst, B. Previtali, S. Deleonibus, Multiple gate devices: advantages and challenges. *Microelectron. Eng.* **80**, 378–385 (2005)
12. P. Banerjee, P. Saha, D.K. Dash, A. Ghosh, S.K. Sarkar, Analytical modeling and performance analysis of graded channel strained dual-material double gate MOSFET, in *4th International Conference on Computing Communication and Automation 2018 (ICCCA)*
13. W. Li, M.D. Brubaker, B.T. Spann, K.A. Bertness, P. Fay, GaN nanowire MOSFET with near-ideal subthreshold slope. *IEEE Electron. Device Lett.* **39**(2), 184–187 (2017). <https://doi.org/10.1109/LED.2017.2785785>
14. R.R. Thakur, P. Singh, Effects of interface charge (Q_{it}) and inter-face trap density (D_{it}) on Al_2O_3 , ZrO_2 and HfO_2 based nano regime multi-gate devices, in *2018 4th International Conference on Devices, Circuits and Systems (ICDCS)*. IEEE, 2018, pp. 58–62
15. R.R. Thakur, P. Singh, Performance reliability of ultra-thin Si-SiO₂, Si-Al₂O₃, Si-ZrO₂ and Si-HfO₂ interface in rectangular steep retrograded nano-regimes devices. *Microelectron. Reliab.* **96**, 21–28 (2019). <https://doi.org/10.1016/j.microrel.2019.02.003>

Performance Analysis of FBMC and OFDM with MIMO for Wireless Communications



Yedukondalu Kamatham and Sushmitha Pollamoni

Abstract In recent years, wireless communication systems are evolved with incredible technological advances which will change the way people communicate and interact. Fifth-generation (5G) wireless communications face a variety of challenges that supports large-scale heterogeneous networks. For supporting 5G communication networks, multicarrier modulations with several transmit and receive antennas have been developed. In this context, this research carries work that deals with comparative analysis of two systems, i.e., multiple-input multiple-output—orthogonal frequency division multiplexing (MIMO-OFDM), filter bank multicarrier (MIMO-FBMC) with different modulation techniques. The power spectral density (PSD) of OFDM, FBMC, MIMO-OFDM, and MIMO-FBMC is estimated and compared. It is found that the spectral efficiency (SE) of FBMC and MIMO-FBMC is more than OFDM and MIMO-OFDM. The bit error rate (BER) and peak-to-average power ratio (PAPR) are also analyzed. The PAPR and BER performance of FBMC is better than OFDM. But the PAPR of MIMO-FBMC is no longer better than MIMO-OFDM. Hence for MIMO wireless communications, OFDM is a better choice, whereas for non-MIMO communications systems FBMC is a good choice.

Keywords MIMO · OFDM · FBMC · PAPR · BER

1 Introduction

In wireless communications (WC), orthogonal frequency division multiplexing (OFDM) is considered as the most privileged technique which dominates the digital broadband communications. The principle behind OFDM is to divide the available spectrum bandwidth into sub-bands with less complex transceiver designs [1]. With many advantages, OFDM suffers from several inadequacies and unsatisfactory requirements for 5GWC. According to 3GPP, the 5G is mostly implemented in massive machine-type communications (mMTC) with the latency of 10 s, ultra

Y. Kamatham (✉) · S. Pollamoni

Department of Electronics and Communication Engineering, CVR College of Engineering, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

521

S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,

Lecture Notes in Electrical Engineering 756,

https://doi.org/10.1007/978-981-16-0749-3_40

reliable, and low latency communication (URLCC) and enhanced mobile broadband (eMBB) with high capacity (500 kmph) of 4 ms latency [2, 3]. For satisfying the future requirements, several modulation candidates are competing in the physical layer to prove the best favorable performer for the next generation of communications systems. The contenders are filter bank multicarrier (FBMC), OFDM, universal filtered multicarrier (UFMC), and generalized frequency division multiplexing (GFDM) [4]. For deciding, whether the candidate is insignificant or not, the following key features are considered and every feature is taken into account with various weighing factors for overall evaluation and decisions. The key possessions are peak-to-average power ratio (PAPR), power spectral density (PSD), spectral efficiency (SE), the design complexity of transceiver, and multiple access interference (MAI) [5].

Multicarrier modulation (MCM) schemes are used for providing SE. OFDM, FBMC, UFMC, and GFDM are examples of MCM systems. The mandate for gigantic data rates and traffic density is enhancing day to day. To gain the attention of this multiple-input multiple-output (MIMO) systems are developed and proven that MIMO is better for problem-solving of traffic capacity in wireless communications [6]. The MIMO is used for multipath transmission with multiple transceiver antennas.

In OFDM, many subcarriers are orthogonal to each other such that each subcarrier overlaps without interference and a guard band is not required for separation of subcarriers. The effect of multipath is encountered by the addition of cyclic prefix (CP) to OFDM symbol and circular convolution takes place which eliminates the inter symbol interference (ISI) due to multipath fading [7]. Due to the addition of the CP, some amount of data is added for each OFDM symbol which leads to a reduction in SE. OFDM also suffers from high PAPR which leads to high power amplifier to operate in the nonlinear region and increases harmonic distortion with out-of-band radiation (OBR) (increase in the spectrum) and in-band radiation (IBR). The harmonic distortion also leads to a reduction in SE. To overcome this problem in OFDM-MCM, an alternate technique is introduced, i.e., FBMC, which is used to provide more SE and to maintain a high data rate. In FBMC, the data of each subcarrier is shaped with the use of a specified filter of well-localized in both time and frequency domain, side lobes also reduced by this reshaping when compared with OFDM. OBR is also controlled and usage of cyclic prefix is not mandatory in FBMC [8] such that more data can be transmitted.

This research work deals with the comparison of SE of FBMC and OFDM. The FBMC and OFDM are combined with MIMO to increase the capacity (or) data rate of the communications with less system complexity. For the estimation of bit error rate (BER), the minimum Ellucian distance method is used. The PAPR and BER are analyzed for both OFDM and FBMC.

The remainder of the paper is followed as: Section 2 deals with system models of OFDM and FBMC. Section 3 presents a brief comparison of OFDM and FBMC with simulation results. Section 4 provides a comparison of MIMO-OFDM and MIMO-FBMC with simulation results. Section 5 describes the discussion on results obtained in Sects. 3 and 4 and in Sect. 6 conclusions are presented.

2 A System Model for OFDM and FBMC

2.1 OFDM

In the OFDM system, the serial bit stream is modulated with quadrature amplitude modulation (QAM) to form N parallel bit streams such that i th parallel bitstream is represented as $s_i(t)$.

$$s_i(t) = a_i g(t) \cos(2\pi f_i t) b_i g(t) \sin(2\pi f_i t) \quad (1)$$

Then $D_i = a_i + b_i$ represents i th constellation symbols and $f_i = f_c + \Delta f$ is the carrier frequency of the i th subcarrier, where a_i represents a real part of complex symbols, b_i is the imaginary part of complex symbols, and Δf is the frequency separation between two subcarriers. The OFDM symbol is represented as

$$s(t) = \sum_{i=0}^{N-1} s_i(t) \quad (2)$$

where the complex envelope of the OFDM system is represented as

$$d(t) = \sum_{i=0}^{N-1} D_i g(t) e^{j2\pi \Delta f t} \quad (3)$$

The $d(m)$ sampling sequence of $d(t)$ is obtained from a period of 0 to T [8, 9].

$$d(m) = \sum_{i=0}^{N-1} D_i g(t) e^{j2\pi \Delta f \cdot \frac{mT}{N}} \quad (4)$$

By the above equation, it proves that the OFDM symbol $d(m)$ is utilizing the inverse fast Fourier transform (IFFT). The OFDM transmitter structure is shown in Fig. 1.

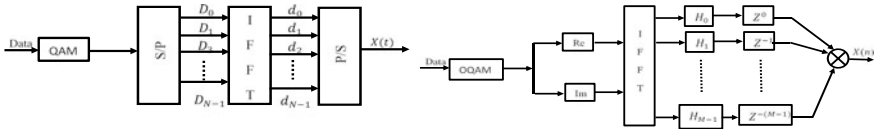


Fig. 1 Transmitter of OFDM and FBMC

2.2 FBMC

In the FBMC, the complex signal is modulated with an offset QAM (OQAM). Whether the complex signal with the real or imaginary part is interleaved, a double sequence length appears.

$$X_i = a_i = D_{i-\text{real}} \quad \text{and} \quad X_{i+1} = b_i = D_{i-\text{imag}} \tag{5}$$

In FBMC, the PHYDYAS [8] filter is used with an overlapping factor of $K = 2, 3, 4$. Such that the frequency response of the filter with M subcarriers is represented as

$$h_{n+1} = 1 - 2H_1 \cos\left(\frac{2\pi n}{KM}\right) + 2H_2 \cos\left(\frac{2\pi 2n}{KM}\right) - 2H_3 \cos\left(\frac{2\pi 3n}{KM}\right) \tag{6}$$

$$1 < n < KM - 1$$

The filter impulse response is with length $L = M * K$ and the sequence of coefficients h_i , i.e., into L length sequence, the data is filtered.

$$y(n) = \sum_{i=0}^{L-1} h_i x(n - i) \tag{7}$$

where $y(n)$ is filtered data in one group. In the frequency domain, the impulse response is represented as

$$H(f) = \sum_{i=0}^{L-1} h_i e^{-j2\pi i f}$$

The filter impulse response is represented as

$$H(Z) = \sum_{i=0}^{M-1} H_p Z^{-p} (Z^M) \tag{8}$$

where

$$H_p(Z^M) = \sum_{k=0}^{K-1} h_{kM+p} Z^{-kM} \quad (9)$$

3 Comparison of OFDM and FBMC Simulation Results

3.1 OFDM

In OFDM, the input bitstream is modulated with the QAM technique such that each subcarrier is passed through IFFT with the orthogonality.

A CP is added and PSD of OFDM is estimated with different fast Fourier transform (FFT) sizes, i.e., FFT = 128, 256, 512, 1024. The PSD is calculated as

$$S_{xx}(\omega) = \lim_{T \rightarrow \infty} E \left[|\hat{x}(\omega)|^2 \right] \quad (10)$$

where

$$\hat{x}(\omega) = \frac{1}{\sqrt{T}} \int_0^T x(t) e^{-i\omega t} dt \quad (11)$$

$x(t)$ is a signal with frequency content, $\hat{x}(\omega)$ is a signal with Fourier transform, $S_{xx}(\omega)$ is the PSD of the required signal. The PSD of OFDM with different FFT sizes is shown in Fig. 2 where spectral density (SD) defines the signal strength and successful transmission of bits in specified bandwidth over a time period. A modulated SD is efficient when the strength (middle part) is closer to the normalized frequency. For FFT size = 1024, the SE is good. Hence, higher FFT sizes are preferred for OFDM.

3.2 FBMC

FBMC is one of the MCM techniques. The input bitstream of FBMC is modulated with the OQAM technique such that each OQAM symbol is passed through FFT. Each FFT symbol is filtered with PHYDYAS filter with coefficients $K = 2, 3, 4$. For different values of K and various FFT sizes = 512 and 1024, the PSD of the FBMC system is shown in Fig. 3, and the comparison of FBMC is estimated with different values of filter coefficients K . By increasing the filter coefficient value, the

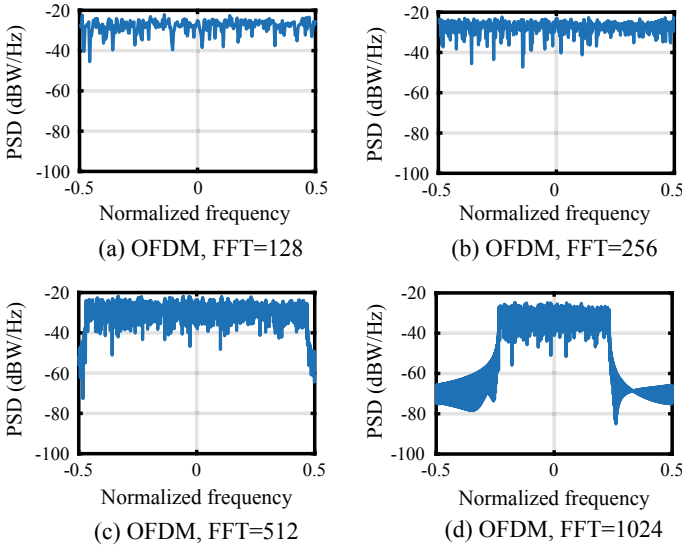


Fig. 2 PSD of OFDM with different number of FFT sizes

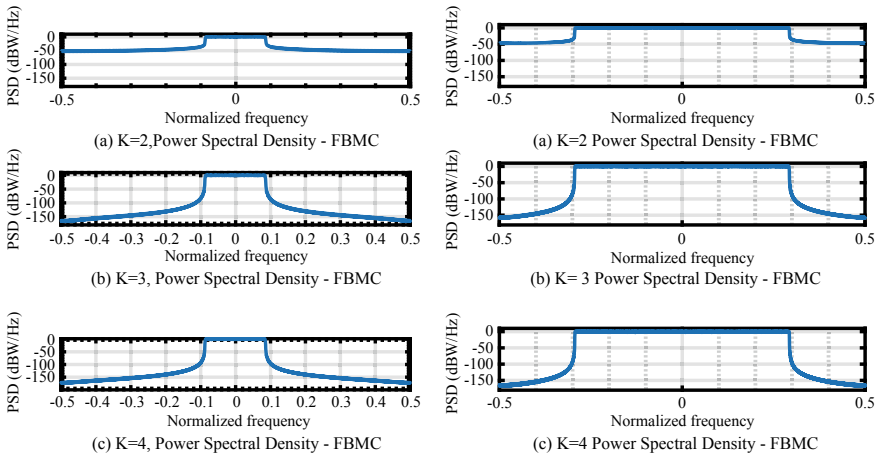
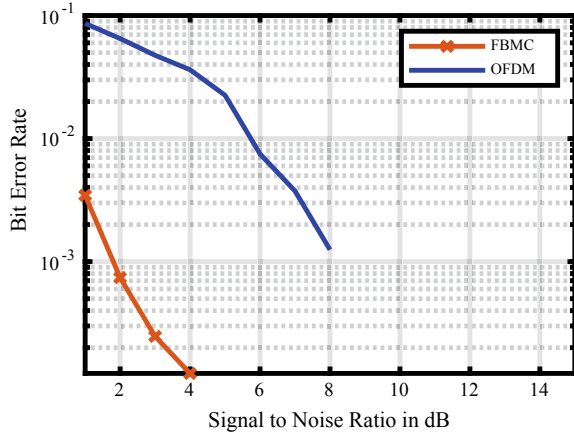


Fig. 3 PSD of FBMC is compared with different k values and FFT sizes

PSD parameter in the y-axis gets decreases to the center of normalized frequency at $K = 4$, the spectrum is more difficult with the value $(-180 \text{ to } 180 \text{ dBW/Hz})$ of FFT size = 512 (Fig. 3). With increasing the FFT sizes, such that FFT size = 512 is good for better filter design and SE.

By comparing Figs. 2 and 3, the PSD of FBMC is less than OFDM. When compared with FFT size = 1024 in Figs. 2 and 3 the filter coefficients $K = 3$ and 4, the PSD is less in FBMC, i.e., the bandwidth efficiency is more in FBMC than

Fig. 4 Comparison BER of OFDM and FBMC



OFDM. The SE in OFDM decreases by the addition of side lobes which occupies more bandwidth. The addition of side lobes is eliminated in FBMC with the help of respective filters and usage of cyclic prefix is not mandatory. Therefore, the SE is more in FBMC than OFDM.

1. **Bit error rate (BER) performance:** BER is defined as the number of errors of bits occurred per unit time. BER is a function of $\frac{E_b}{N_0}$ and represented as

$$BER = \frac{1}{2} \operatorname{erfc} \left(\sqrt{\frac{E_b}{N_0}} \right) \tag{12}$$

The BER of OFDM and FBMC is shown in Fig. 4. With increase in SNR, the BER performance is decreased. By comparing, the BER performance of OFDM is higher than the FBMC such that BER performance is better in FBMC than OFDM.

2. **PAPR analysis:** PAPR is defined as the ratio of instantaneous power to the signal average power.

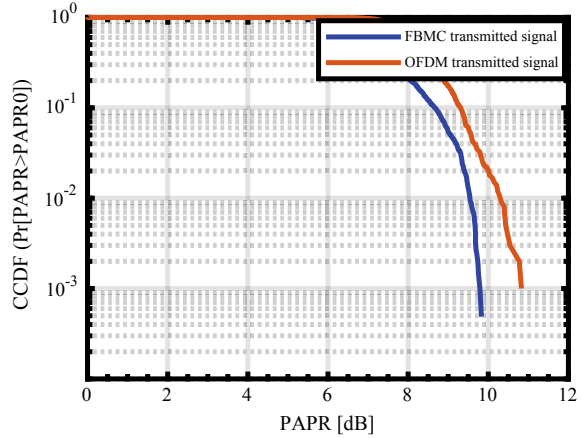
$$PAPR[x_n] = 10 \log_{10} \frac{\max |x_n|^2}{E[|x_n|^2]} \tag{13}$$

where $|x_n|$ is the signal peak power. The PAPR is evaluated by cumulative distributed function (CDF) within the threshold level.

$$F_z(z) = 1 - e^{-z} \tag{14}$$

The PAPR beyond the threshold level is expressed in terms of complementary cumulative distributed function (CCDF($\tilde{F}_z(z)$)) as $\tilde{F}_z(z) = 1 - (F_z(z))^n$.

Fig. 5 PAPR comparison of OFDM and FBMC



The estimation of PAPR performance for both OFDM and FBMC is shown in Fig. 5. At $CCDF = 10^{-2}$, the PAPR of FBMC is 9.5 dB and PAPR of OFDM 10.5 dB. The PAPR of FBMC is less than OFDM.

4 Comparison of MIMO-OFDM and MIMO-FBMC Simulation Results

4.1 MIMO

In MIMO systems, the multiple streams of data are transmitted through multiple antennas. These multiple streams are passed through a channel matrix which consists of $N_T N_R$ paths between N_T transmit antennas and N_R receive antennas. The receiver receives the signal vectors by multiple receivers and decodes the signal into the original information. A MIMO system is modeled as $Y = HX + n$.

Then Y is a received signal vector with a length of $Y = [y_1, y_2, \dots, y_N]$ and X is a transmitted signal vector with a length of $X = [x_1, x_2, \dots, x_N]$, H is a channel matrix with a length of $M \times N$ matrix and n is the noise vector.

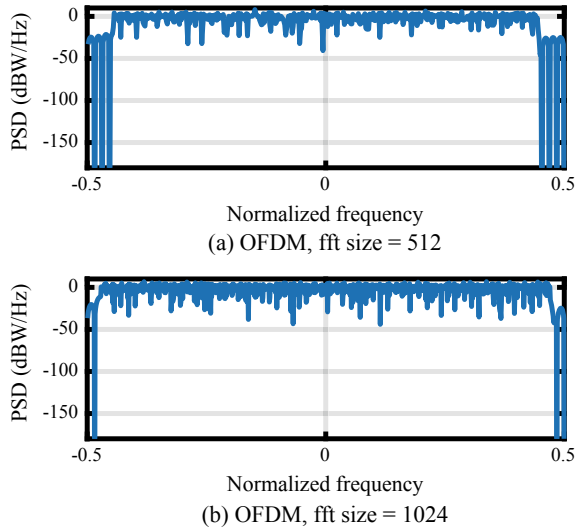
1. MIMO-OFDM

To increase the high data rate, the OFDM signal is combined with MIMO of 2×2 such that PSD of MIMO-OFDM with different FFT sizes is shown in Fig. 6. With the increase in the FFT sizes the spectral density of MIMO-OFDM increases.

2. MIMO-FBMC

To increase the SE, the 2×2 MIMO system is combined with FBMC, and PSD of MIMO-FBMC is estimated with different values of K as shown in Fig. 8. With

Fig. 6 MIMO-OFDM with FFT size 512 and 1024



an increase in K value at FFT size = 512, 1024, the spectral density of FBMC also decreases. The SE decreases with an increase in FFT sizes rather than an increase in different K values. The SE for FFT = 512 is more efficient than FFT = 1024. By comparing Figs. 6 and 7 at FFT sizes = 512, 1024 the PSD of MIMO-FBMC is more. With an increase in K value, the spectral density decreases. The SE is more in MIMO-FBMC rather than in MIMO-OFDM because of the complex filter nature of the FBMC transmitter.

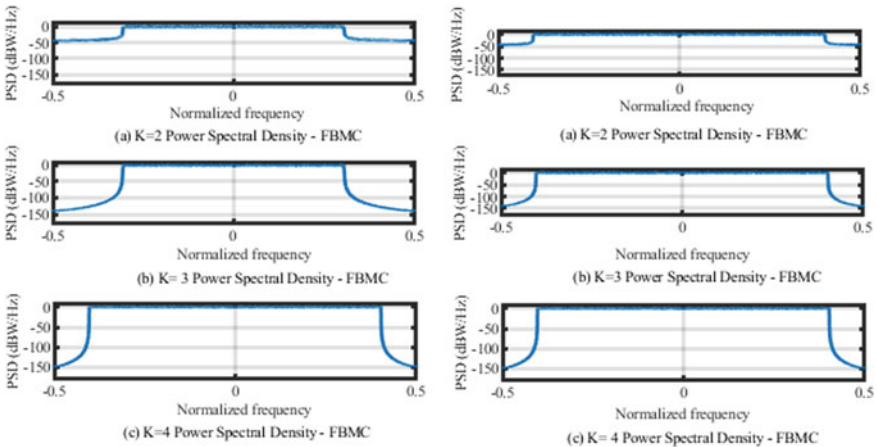
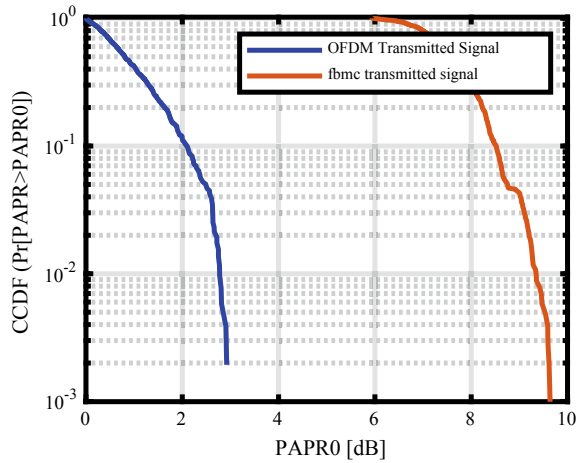


Fig. 7 Comparison of MIMO-FBMC with different k values and FFT sizes

Fig. 8 Comparison of PAPR for MIMO-OFDM and MIMO-FBMC



In Fig. 8, the comparison of PAPR performance for MIMO-OFDM and MIMO-FBMC. According to the comparison, the PAPR of MIMO-OFDM is less than the MIMO-FBMC.

5 Discussion on Simulation Results

OFDM system has been developed to decrease the ISI due to multipath fading. Reduction in SE is one of the major drawbacks of the OFDM system. To overcome the reduction of SE, the FBMC system has been developed. In FBMC, instead of using a cyclic prefix, the filter parameters are added such that side lobes of the signal decrease. The BER versus SNR is given in Table 1. BER of OFDM and FBMC is estimated and is less than the OFDM. The PAPR of OFDM, FBMC, MIMO-OFDM, and MIMO-FBMC is explained in detail in Sect. 3 such that the PAPR of FBMC is less when compared to OFDM. The PAPR of MIMO-OFDM is less compared with

Table 1 SNR versus BER

SNR	BER			
	OFDM BER	FBMC BER with varying K values		
		$K = 2$	$K = 3$	$K = 4$
1	0.08625	0.0035511	0.0036267	0.003582
2	0.065	0.0015388	0.0013298	0.00061759
3	0.0475	0.0008286	0.00036267	0.00024704
4	0.03625	0.000011837	0.00024178	0.00012352
5	0.0225	0	0	0

MIMO-FBMC. For the MIMO system, the PAPR of OFDM is better than FBMC according to the literature.

6 Conclusions

In this paper, the SE and PAPR of OFDM and FBMC are compared. Because of the addition of CP in OFDM, the SE gets reduced and due to high PAPR, the HPA enters nonlinear regions and leads to harmonics distortions includes OBR and IBR. To overcome these disadvantages, FBMC is developed. To gain spatial transmission, the MIMO system is combined with both OFDM and FBMC. The SE of MIMO-OFDM and MIMO-FBMC is estimated and compared such that SE for both FBMC and MIMO-FBMC is better than OFDM and MIMO-OFDM. The PAPR and BER of FBMC are better than OFDM, but PAPR of MIMO-FBMC is not better than MIMO-OFDM and also the transceiver design complexity is more in FBMC. Hence, in MIMO systems, OFDM can be used with an appropriate choice of companding techniques reported in the literature. But in non-MIMO wireless systems, FBMC is a better choice for PAPR and BER performance.

Acknowledgements This work has been carried out under the project entitled “Study and Implementation of Self-Organized Femtocells for Broadband Services to Indoor Users in Heterogeneous Environment” sponsored by AICTE, New Delhi under Research Promotion Scheme (RPS), Vide sanction letter No: 8-30/RFID/RPS/POLICY-1/2016-2017, Dated: 02.08.2017.

References

1. C. Bockelmann et al., Massive machine-type communications in 5G: physical and MAC-layer solutions. *IEEE Commun. Mag.* **54**(9), 59–65 (2016)
2. ITU-R Rec. ITU-R M. 2083-0, IMT Vision-Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond (2015)
3. P. Popovski et al., Wireless access in ultra-reliable low-latency communication (URLLC). *IEEE Trans. Commun.* **67**(8), 5783–5801 (2019)
4. R. Gerzaguët et al., The 5G candidate waveform race: a comparison of complexity and performance. *EURASIP J. Wirel. Commun. Netw.* **2017**(1), 1–14 (2017)
5. A. Husam, Z. Kollar, Complexity comparison of filter bank multicarrier transmitter schemes, in *2018, 11th International Symposium of Communication Systems, Networks and Digital Signal Processing (CSNDSP)* (2018)
6. R. Zakaria, D. Le Ruyet, A novel filter bank multicarrier scheme to mitigate the intrinsic interference: application to MIMO systems. *IEEE Trans. Wireless Commun.* **11**(3), 1112–1123 (2012)
7. I. Das, R.N. Shaw, S. Das, Location-based and multipath routing performance analysis for energy consumption in wireless sensor networks, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh, vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_59

8. A. Ghassemi, T. Gulliver, Intercarrier interference reduction in OFDM systems using low complexity selective mapping. *IEEE Trans. Commun.* **57**(6), 1608–1611 (2009)
9. S. Shreyan, S. Rajarajan, Ch. Abhijit, Environment-adaptive concurrent companding and bias control for efficient power amplifier operation. *IEEE Trans. Circuits* **58**(3), 607–618 (2010)

Magnetization Pattern Study of Unit Domain Multiferroic Nanomagnet for Spintronics Devices



Amlan Sen, Rabindra Nath Shaw, and Ankush Ghosh

Abstract Nanomagnets are used in spintronics/straintronics devices, with their non-volatile and non-leakage feature, and are opening new possibilities of replacing conventional digital logic using transistor-based electronics. Correlated switching of spins is presented in nanomagnets, and this novel feature offers credibility of replacing CMOS VLSI technology with high integration density and energy efficiency. In the present paper, the authors studied the unit domain cell of a circular-shaped nanomagnet cylinder using micromagnetic simulation software OOMMF. The simulation result would help to determine the minimum domain size in nanomagnet-based devices using spintronics.

Keywords Spintronics devices · Nano-magnetic devices · Multiferroic · OOMMF

1 Introduction

Over the last 70 years, the complementary metal oxide semiconductor (CMOS) device technology has dominated the field of electronics. The CMOS-based VLSI chips shrink rapidly at an incredible pace, according to Moor's prediction. This continued downscaling of device size may reach at such size within a few years that quantum mechanical, and other effects will be significantly high. As a result, heat generation by SOC circuits is tremendously high. This rapid increase in generated heat is unmanageable as available heat sinks are unable to dissipate such amount of heat quickly from the circuit.

Another need of future electronic devices is that it can operate with a few order of lower magnitude of energy dissipation than current CMOS device. It is preferable that handheld electronic gadgets and medical equipments (e.g. pacemaker) would

A. Sen · A. Ghosh (✉)

School of Engineering and Applied Sciences, The Neotia University, Sarisha, West Bengal, India

R. N. Shaw

Department of Electrical, Electronics and Communication Engineering, Galgotias University, Greater Noida, India

consume power as little as possible, and some cases it can harvest energy from the ambient without requiring a separate power supply.

To overcome the above-mentioned issues, researchers have to find out some entire new technology, which can sort out all the issues together. One of the possible solutions is spintronics/straintronics [1–6] where electron spins are used for storing data. This technology does not depend upon the use of charge and control of physical movement of electrons. Therefore, it produces lesser amount of heat and requires very little power to operate. Spintronics is a technology which exploits the spin degree of freedom of electrons and its associated magnetic moment, rather than its electronic charge. The manipulation of magnetic moment of electrons is achieved by transfer of spin angular momentum from spin polarized current [7–9]. The spintronic devices can be built using nanomagnets. Therefore, it is possible to encode logic bits processing and storing information by nanomagnets due to their non-volatile and non-leakage property [10–14].

One of the intrinsic property of ferromagnetic materials is magnetic anisotropy, which refers to the tendency of alignment of magnetization of ferromagnetic materials along certain crystallographic direction. The crystallographic axes, along which the magnetization tends to align, are called easy axes and other axes, along which it is very difficult to saturate the magnetization, are called hard axes. The shape of the hysteresis loop is a perfect rectangle when the direction of magnetization is parallel to the easy axis and the coercive force depends on saturation magnetization and anisotropy. In ferromagnetic materials, individual domain can be magnetized very easily, and therefore, small magnetic field is needed for it. It can be assumed with reasonable accuracy that all individual domains are magnetically saturated even the absence of external field, but in the total specimen net magnetization is zero.

With the application of external magnetic field, domain walls start moving resulting in the growth of the domains which are already oriented in the direction of field. This process requires little energy. If the applied magnetic field is large enough to overcome the anisotropy energy, direction of domain magnetization will be toward that easy axis which is nearest to the direction of applied field. But in the materials with high anisotropy energy, both movement of domain walls and change the orientation of magnetization are difficult. Such materials are difficult to put in saturation or demagnetize totally. Impurities, defects restrict the movement of domain wall and thus increase coercivity. A hysteresis loop having large area implies large coercivity and the area inside the loop represents energy dissipation per cycle of magnetization.

Magnetocrystalline anisotropy originates from spin-orbit coupling. The electronic orbits depend upon the crystallographic axes and the spin-orbit coupling forces the electronic spins to be aligned along certain crystallographic axes. Usually, anisotropy energy is expressed as a function of the power of trigonometric functions of angle made by magnetization with the easy axis. It plays an important role for the application in magnetic memories, loudspeakers, electric motors, microphones, etc. Magnetic anisotropy not only varies from material to material, but it can also be changed on the manipulation of the shape of nanomagnets.

The exchange interaction between two nearest neighbor spins is roughly 10^3 times larger than the dipolar interaction between them. However, the exchange interaction

is a short-range interaction (decays exponentially with distance), whereas dipolar interaction is a long range interaction. If in a material contest exists between exchange and dipolar interactions, magnetic domains are formed and this usually occurred in large sample. Within a magnetic domain, the exchange interaction dominates and the spins are aligned parallel along a certain direction. However, for a ferromagnet with dimension greater than the dimension of a domain, the long-range dipolar interaction dominates over the exchange interaction. This leads to the formation of a number of domains separated by domain walls. Magnetic domains reduce the magnetostatic stray field energy associated with the uncompensated free charges on the surface of ferromagnets.

Generally, ferromagnetic materials are made up of several domains. But under certain conditions, they can be single domain also. If a ferromagnet is saturated by applying a large enough magnetic field, the material becomes single domain. If the dimension of the material is small enough, the short-range exchange interaction dominates over the demagnetizing field even in the absence of an external magnetic field. Then also the element can be in a single domain state.

The domains are separated by domain walls where the direction of magnetic moment changes gradually to minimize the total energy associated with the domain wall. Then calculated thickness of domain wall from a simple model can be expressed as [15–17]

$$\delta = \sqrt{\frac{J_e \pi^2 S^2}{K a}}$$

where J_e , K and a are the exchange integral, anisotropy constant, and lattice constant, respectively. A variety of domain walls are observed in magnetic materials, out of which Bloch wall and Néel wall are very common. When the magnetization from one domain to its neighboring domain rotates along the axis normal to the domain wall, it is called Bloch wall. But in case of Néel wall, the rotation axis is parallel to the domain wall. The typical width of domain wall for Co, Ni, and permalloy are about 10, 60, and 57 nm [18, 19].

The basic structure of magnetic multilayers composed of alternate ferromagnetic and nonmagnetic metals form a stack. The thickness of each layer ranges between 1 and 10 nm, and number of layers varies between 3 and 100. Fe, Co, Ni, and their alloys are usually chosen for ferromagnetic materials while for choosing nonmagnetic metals, Cr, Ru or noble metals Cu, Ag, and Au are used [20].

In the present work, we studied energy states for a circular-shaped cylinder. It was already observed through micromagnetic simulation that diamond and concave-shaped nanomagnets entail coherent switching modes with an incoherent percentage rate that is near to zero, for various thicknesses [21]. With increasing thickness, concave-shaped nanomagnets are coherent and exhibits reliable magnetization states while the diamond-shaped nanomagnets shows vorticity in magnetization states and thus prone to incoherent [21].

To study the magnetization switching in our mode, object-oriented micromagnetic framework (OOMMF) software [22] has been used. In order to study the magnetization dynamics in nanomagnets, various energy terms (e.g., Zeeman, Exchange) must be considered.

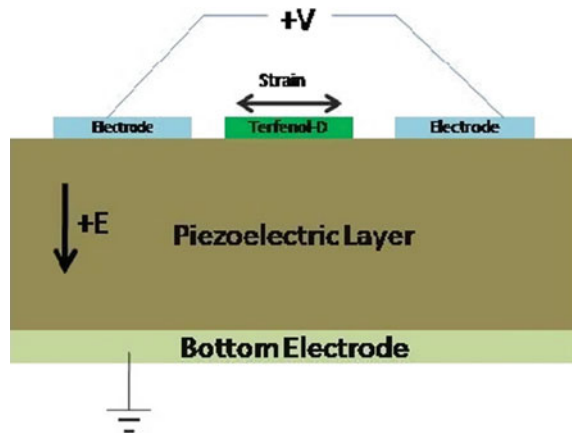
2 Model

There exist different types of interactions in micromagnetics, which mainly varies with the wide range in length scale. In short range, exchange interaction dominates, whereas dipolar interaction occurs in long range. Also, the magnetic state in a material depends on the length scale. One of the smallest features in vortex structure is the region where orientation of magnetization is perpendicular to the surface. In Neel wall, one of the characteristic features is domain wall width, which can be defined by $\sqrt{A/K_m}$ [23], where K_m is the stray field energy constant and K_m is $\frac{1}{2}\mu_0 M_s^2$. For Bloch wall, the domain wall width is $\sqrt{A/K}$ [23], where K is the anisotropy constant. Both Bloch wall and Neel wall width are called exchange length, and in order to get better simulation result in micromagnetic calculations, it should not be kept above to the minimum values of the two exchange lengths.

In this paper, we investigated the straintronics device which is a cylindrical circle in shape as shown in Fig. 1. In our study, for choosing magnetostrictive material, we select Terfenol-D for its high magnetostrictive value, and therefore, switching of magnetization states needed very little power. To study the minimum domain size of straintronics devices, we have created flower and vortex magnetization state in the cylindrical circle and after relaxation, studied its energy dissipation at different single domain length.

To study the energy dissipation at two different magnetization states in our model, we consider various energy terms used by OOMMF. The total energy of a nanomagnet

Fig. 1 Cross-sectional schematic of a multiferroic nanomagnet undergoing strain



of volume Ω can be defined by Gibbs free energy equation:

$$U_i = \oint \left\{ \begin{array}{l} A \left[(\nabla m_x)^2 + (\nabla m_y)^2 + (\nabla m_z)^2 \right] - \frac{1}{2} \mu_0 \vec{H}_d \cdot \vec{M} \\ + \left[B_1 (\alpha_1^2 *_{xx} + \alpha_2^2 *_{yy} + \alpha_3^2 *_{zz}) \right. \\ \left. + B_2 (\alpha_1 \alpha_2 *_{xy} + \alpha_2 \alpha_3 *_{yz} + \alpha_3 \alpha_1 *_{xz}) \right] - \mu_0 \vec{M} \cdot \vec{H} \end{array} \right\} d\Omega \quad (1)$$

In Eq. (1), the four terms represent exchange energy, magnetostatic energy, magnetoelastic energy, and Zeeman energy, respectively. In this equation, A represents exchange constant, B_i is magnetoelastic coupling constants, and direction cosines, α_i , under the strain $*_{ij}$. Also, it is assumed that the nanomagnet has random polycrystalline orientation, and therefore, magnetocrystalline anisotropy is neglected.

When a nanomagnet is placed under the influence of external magnetic field \vec{H}_{eff} , the magnetization dynamics can be described by the Landau–Lifshitz–Gilbert (LLG) equation [24, 25]:

$$\frac{d\vec{M}(t)}{dt} = -\gamma \vec{M}_i(t) \times \vec{H}_{\text{eff}}^i(t) - \frac{\alpha\gamma}{M_S} \left[\vec{M}_i(t) \times \left(\vec{M}_i(t) \times \vec{H}_{\text{eff}}^i(t) \right) \right] \quad (2)$$

In this equation, γ represents gyromagnetic ratio, M_S is the saturation magnetization and α is the dimensionless Gilbert damping parameter. Therefore,

$$\vec{H}_{\text{eff}}^i(t) = -\frac{1}{\mu_0 \Omega} \frac{\partial U_i(t)}{\partial \vec{M}_i(t)} = -\frac{1}{\mu_0 \Omega M_S} \nabla_{\vec{m}} U_i(t) \quad (3)$$

To calculate stress anisotropy energy, y-axis is chosen for application of stress. Therefore, stress anisotropy energy for volume Ω and magnetostrictive coefficient λ_S of the nanomagnet is

$$U_{\text{Stress - Anisotropy}}(t) = -\left(\frac{3}{2} \lambda_S \right) \sigma(t) \Omega m_y^2(t) \quad (4)$$

The effective field produces by the stress anisotropy (Eq. 4) is:

$$H_{\text{eff} - \sigma}(t) = \left(\frac{3}{\mu_0 M_S} \lambda_S \right) \sigma(t) m_y(t) \quad (5)$$

This effective field, having direction only along the y-axis, is used for OOMMF simulation. To study magnetization pattern in the nanomagnet, OOMMF was used as micromagnetic simulator. This OOMMF simulator used LLG equation and performed integration over time, while include the energy contributions to the exchange interaction, the magnetostatic interaction, the anisotropy, and Zeeman energy. To calculate the minimum domain limit, the discretized cell size is taken as the variable parameter in the simulation and implemented in the Cartesian coordinate

system. For micromagnetic simulation, following parameters related to Terfenol-D are used [26–32]:

Parameters	Values
Exchange stiffness, A	$9 \times 10^{-12} \text{ J m}^{-1}$
Saturation magnetization, M_S	800 KAm^{-1}
Anisotropy constant, K_1	0 J m^{-3}
Gilbert damping constant, α	0.1 [16]
Saturation magnetostriction $3/2\lambda_S$	900×10^{-6}
Young's modulus, Y	80 GPa

3 Results and Discussions

In this paper, we have studied the magnetization pattern and calculate the single domain limit for our proposed circular shaped cylindrical model. For that purpose, vortex (or curling state) and flower states (sometimes this state is also called modified single domain state or splayed state) are created in the nanomagnet, and then the system was relaxed. We used the object-oriented micromagnetic framework (OOMMF) as simulation software for studying the magnetization pattern and for other purposes. Figures 2 and 3 depict the magnetization pattern of two relaxed states.

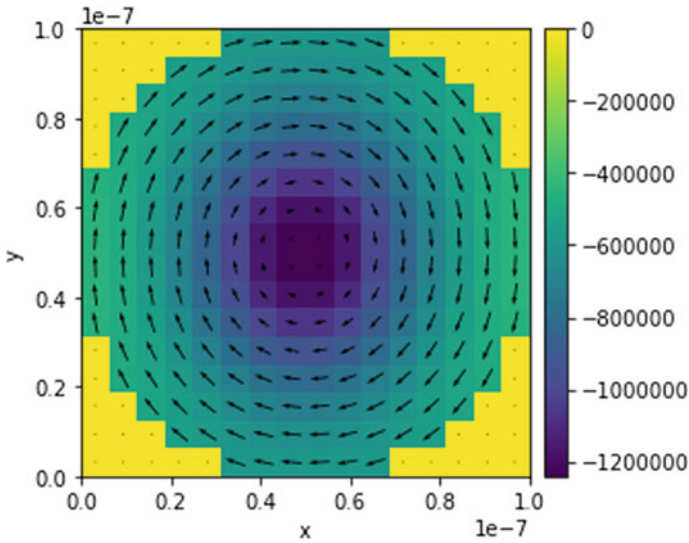


Fig. 2 Magnetization pattern for the vortex state in the circular-shaped nanomagnet

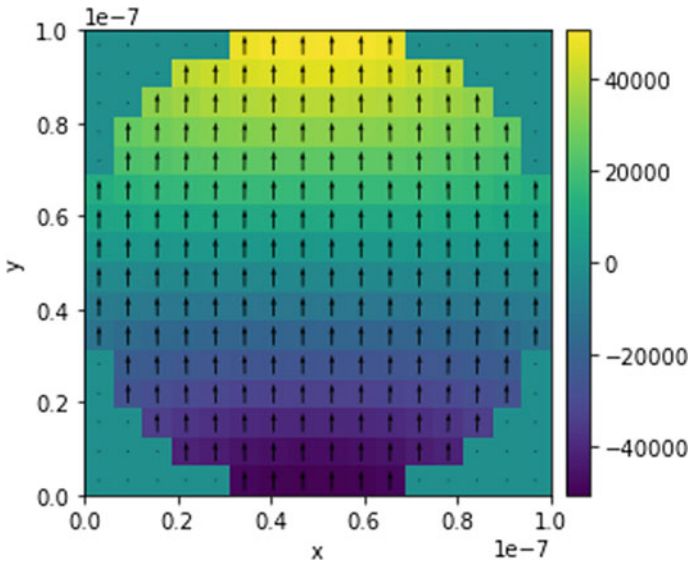


Fig. 3 Magnetization pattern for the flower state in the circular-shaped nanomagnet

Usually, the vortex state is created inside a sample during switching from both parallel to antiparallel state of magnetization between two layers and vice versa. This non-uniform reversal of magnetization was predicted by Lee et al. [33–36], which involves vortex formation and annihilation.

Nano-sized magnetic materials have interesting magnetic properties which make them different from their bulk. To explain the behaviors of nano-magnetic materials, Maxwell’s equations are insufficient but by including quantum mechanical effects, the theory of micromagnetics can do this successfully. With the help of micromagnetics theory, formation of different states such as vortex, leaf and flower states can be analyzed and explained easily. This theory is widely applied in modern day hard disk and other storage devices of digital data.

The LLG equation is an ordinary differential equation in time which includes the exchange interaction, dipolar interaction, magnetocrystalline anisotropy, and the Zeeman field. This equation can be solved by using standard ODE solvers like Runge–Kutta method or Euler method. The LLG equation can be solved numerically with the help of the finite difference method (FDM) or finite element method (FEM). In OOMMF, LLG equation is solved by using FDM method. The initial conditions and input parameters are specified in OOMMF. The initial temperature was set to 0 K for simulation. The sample is divided into square cells, having dimension less than or equal to exchange length. While calculating total energy in each cell, it considers exchange energy, self-magnetostatic energy, magnetocrystalline anisotropy energy, and Zeeman energy. Two types of evolvers are used for updating magnetization configuration, time evolver, and energy minimization evolver. The first one tracks the LLG dynamics, and the second technique calculates the local minima in energy by

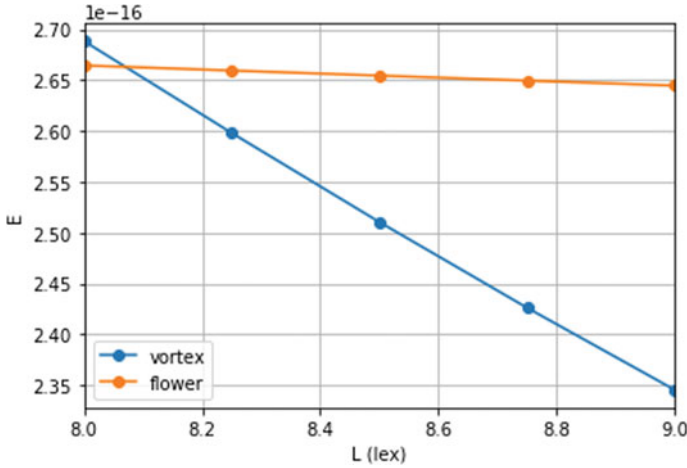


Fig. 4 Energy diagram for the flower and vortex state in the circular-shaped cylinder for varying discrete cell dimension

using energy minimization techniques. The two drivers (time driver and minimization driver) control their respective evolvers.

In order to determine the single domain limit (this is the size L of equal energy for both the flower and vortex state), we have to make a plot of energy dissipation for flower state and vortex state as a function of domain length. Figure 4 depicts that graph. This graph clearly shows that for a large variation in domain length of nanomagnets for our model, energy dissipation is almost constant in case of flower state, while it reduces gradually for vortex state. The two graphs crosses at a point, which gives the value of single domain limit.

4 Conclusions

In this paper, we have observed the effect of size of a single domain on the magnetization states of nanomagnets of straintronic devices. Two different magnetization patterns (vortex state and flower state) were examined in the nanomagnet and then system was relaxed. It was observed that nanomagnets having smaller domain size are less energy dissipative in case of vortex magnetization pattern. But there is a very little effect on energy dissipation with the variation of domain length for flower state. So for smaller nanomagnets, straintronic switching is a good choice and also energy-efficient. Straintronic-based logic devices are less energy dissipative than CMOS-based devices. Furthermore, both types of devices are almost equivalent in terms of propagation delay and switching error probability. But there is a major difference, CMOS devices are volatile, whereas straintronic-based devices are non-volatile, making them suitable for non-volatile memory with high packing density.

While non-volatile devices exhibit both logic and memory capability, it suffers from resiliency, error probability, and energy dissipation due to thermal fluctuation and inter-magnet coupling effect. These problems are still unresolved and need to carry out further research work.

References

1. M.S. Fashamin, K. Roy, J. Atulasimha, S. Bandyopadhyay, *Nanotechnology* **22**, 155201 (2011)
2. M.S. Fashami, J. Atulasimha, S. Bandyopadhyay, *Nanotechnology* **23**, 105201 (2012)
3. A. Imre, G.H. Bernstein, W. Pord, V. Metlushko, *IEEE Trans. Nanotech.* **1**, 209 (2002)
4. A. Sarkar, S. Halim, A. Ghosh, S.K. Sarkar, Implementation of PMN-PT/Ni based NOR gate with biaxial anisotropy off ultra low energy dissipation. *J. Nanoelectron. Optoelectron.* **12**, 1–6 (2017)
5. A. Sarkar, P.K. Dutta, A. Ghosh, S. Roy, S.K. Sarkar, Implementation of universal gates (NAND) based on nanomagnetic logic using multiferroics. *Quantum Matter* **5**, 505–509 (2016)
6. A. Ghosh, A. Basu, T.S. Das, V.H. Mankar, D. Samanta, S.K. Sarkar, Single spin logic realization of robust image watermarking in spatial domain, in *2008 IEEE Region 10 Colloquium and the Third ICIS*, IIT Kharagpur, India, 8–10 Dec 2008
7. E.Y. Tsybal, I. Zutic, *Science—Handbook of Spin Transport and Magnetism* (2016). p. 11
8. A. Sen, A. Ghosh, Performance study of PMNPT/Ni-based multiferroic nanomagnet for straintronics device applications, in *Advances in Intelligent Systems and Computing*, ed. by L.C. Jain, M. Virvou, V. Piuri, V.E. Balas (Springer, Berlin, 2019), pp. 15–24. ISSN: 2194-5357
9. A. Ghosh, S. Sarkar, D. Ray Chaudhuri, S.K. Sarkar, A power efficient binary multiplier circuit with overflow detection using single spin logic circuit: design and implementation. *Adv. Sci. Lett.* **2**, 86–96 (2009)
10. S. Sarkar, S.A. Alam, V. Ravindran, A. Banerjee, A. Debnath, A. Ghosh, Spintronics based systems for AI nano robots, in *International Conference on Computing, Power and Communication Technologies* (IEEE Conference Record No 42736) (2018)
11. F.J. Albert, J.A. Katine, R.A. Buhrman, D.C. Ralph, Spin-polarized current switching of a Co thin film nanomagnet. *Appl. Phys. Lett.* **77**(23), 3809 (2000)
12. K. Roy, S. Bandyopadhyay, J. Atulasimha, Hybrid spintronics and straintronics: a magnetic technology for ultra low energy computing and signal processing. *Appl. Phys. Lett.* **99**(6), 063108 (2011)
13. K. Roy, S. Bandyopadhyay, J. Atulasimha, Energy dissipation and switching delay in stress-induced switching of multiferroic nanomagnets in the presence of thermal fluctuations. *J. Appl. Phys.* **112**(2), 023914 (2012)
14. A. Ghosh, A. Basu, T.S. Das, V.H. Mankar, D. Samanta, S.K. Sarkar, Single spin logic realization of a robust spatial domain image watermarking. *Adv. Sci. Lett.* **2**, 1–11 (2009)
15. T. Shinjo, *Nanomagnetism and Spintronics*, p. 31
16. A. Ghosh, K. Chakraborty, A. Kramakar, S.K. Sarkar, Application of spintronic device for realization of an edge-triggered phase detector, in *International Conference on Communication, Computers and Devices*, Kharagpur, India, 10–12 Dec 2010
17. D. Samanta, A. Basu, T.S. Das, V.H. Mankar, A. Ghosh, M. Das, S.K. Sarkar, SET based logic realization of a robust spatial domain image watermarking, in *ICECE 2008*, Department of EEE, BUET, Dhaka-1000, Bangladesh held on 20–22 Dec 2008
18. J. Silcox, *Philos. Mag.* **8**, 7 (1963)
19. T. Suzuki, *J. Phys. Colloq.* **32**, C1 134 (1971)
20. S. Chikazumi, *Physics of Ferromagnetism* (Oxford University Press, New York, United States, 1997)
21. M. Salehi-Fashami, N. D'Souza, Exploring performance, coherence, and clocking of magnetization in multiferroic four-state nanomagnets

22. M.J. Donahue, D.G. Porter, OOMMF User's Guide, Version 1.0, Interagency Report NISTIR 6376 (1999)
23. R. Schäfer, A. Hubert, *Magnetic Domains—The Analysis of Magnetic Microstructures* (Springer, Verlag, 1998)
24. I.D. Mayergoyz, G. Bertotti, C. Serpico, *Nonlinear Magnetization Dynamics in Nanosystems*, 1st edn. (Elsevier Science, Amsterdam, 2009), p. 480
25. G. Dewar, Effect of the large magnetostriction of Terfenol-D on microwave transmission. *J. Appl. Phys.* **81**(8), 5713 (1997)
26. S. Chikazumi, *Physics of Magnetism* (Wiley, New York, 1964).
27. K. Ried, M. Schnell, F. Schatz, M. Hirscher, B. Ludescher, W. Sigle, H. Kronmüller, *Phys. Status Solidi A* 167–195 (1998); R. Abbundi, A.E. Clark, *IEEE Trans. Mag.* **13**, 1547 (1977)
28. R.A. Kellogg, A.B. Flatau, *J. Intell. Mater. Syst. Struct.* **19**, 583 (2008)
29. J. Walowski, M. Djorjevic Kaufman, B. Lenk, C. Hamann, J. McCord, M. Münzenberg, *J. Phys. D Appl. Phys.* **41**, 164016 (2008)
30. S.K. Sarkar, T. Datta, A. Ghosh, J.C. Singh, G. Prakash, D. Samanta, Efficient nanodevices for advanced ultradense VLSI systems, in *International Conference on Modelling and Simulation (MS07)* held on 3–5 Dec 2007 at University of Calcutta, pp. 775–777
31. A. Ghosh, J. Gope, T. Datta, B. Roy, S.K. Sarkar, Spintronic device based power efficient VLSI chip design for universal code converter. *Can. J. Pure Appl. Sci.* **2**(3), 469–474 (2008)
32. A. Sen, A. Ghosh, Study of single domain limit of a multiferroic nanomagnet for spintronics devices, in International Seminar Cum Research Colloquium on MEMS Based Sensors and Smart Nanostructured Devices on 27 Dec 2019 at Jadavpur University
33. A. Sarkar, S. Halim, A. Ghosh, S. Roy, Straintronics-MTJ based 3-input universal logic gate (NOR) with high efficiency, in *International Conference on Computational Science and Engineering [ICCSE2016]*, 4th–6th Oct 2016 organized by RCC Institute of Information Technology, Kolkata
34. A. Sarkar, S. Halim, A. Ghosh, S. Roy, Spintronics device based low power RFID system: design and implementation, in *International Conference on Microelectronics, Computing and Communication (MicroCom 2016)*, organized by National Institute of Technology, Durgapur, held on 23–25 Jan 2016
35. N. Basanta Singh, S. Deb, A. Ghosh, M.A. Goutham, S.K. Sarkar, Single electron threshold logic based digital computer, in *15th International Workshop on the Physics of Semiconductor Devices (IWPSD)* held on 15–20 Dec 2009 at Jamia Millia Islamia, New Delhi
36. K.J. Lee, A. Deac, O. Redon, J.P. Nozières, B. Dieny, *Nat. Mater.* **3**, 877 (2004)

Advanced RSSI-Based Wi-Fi Access Point Localization Using Smartphone



Debajyoti Biswas, Suvankar Barai, and Buddhadeb Sau

Abstract The geographical position of sensors is important information in wireless sensor networks (WSNs). The range-based localization systems by received signal strength indicator (RSSI) achieve high accuracy compared to a broad range of other localization approaches. However, RSSI suffers from various environmental dynamic and static catalyst factors. To manage these issues, several methods have been proposed in the literature. It is seen that the existing methods might not perform adequately well in ever-changing environments. Because RSSIs are fluctuating in significant ranges from time to time. For improving the performance, this paper has introduced the scheme of curve fitting-based algorithm, utilizing the most probable RSSIs. Experiments were conducted in the indoor and outdoor, with ten wireless fidelity (Wi-Fi) routers and a smartphone, where routers worked as an access point (AP), and smartphone worked as a station (STA). A Wi-Fi analyzer application (APP) was installed on a smartphone to observe the RSSIs of every router simultaneously.

Keywords Localization · WSNs · Catalyst factors · Curve fitting · Most probable RSSIs · Wi-Fi analyzer APP

1 Introduction

From last few years Wi-Fi positioning system becomes more and more popular due to its several applications and widespread adoption. Among all the positioning techniques, the RSSI has become mainstream with advantages of simple process, low cost, extensive coverage, and no need for additional hardware [1–4], although RSSI-based information commonly affected by various catalyst factors such as reflections, refractions, multipath, moisture, human body, temperature, metals, electromagnetic waves, and electronic devices. [5, 6]. It is impossible to know about a number of catalyst factors that are present in the environment, and how they affect the RSSI-based measurement [7, 8]. However, many algorithms are introduced to achieve higher

D. Biswas (✉) · S. Barai · B. Sau
Department of Mathematics, Jadavpur University, Kolkata, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_42

543

accuracy. Xue et al. [9] proposed an algorithm by using the average of maximum RSSIs and they have compared their algorithm with other existing algorithms under both static and dynamic conditions [10]. However, our experiment results show that the largest RSSIs give huge errors compared to mode, average, and median values. Also, their proposed signal attenuation model produces enormous errors for measurement. In this work, routers are placed in multiple locations, and from different positions, the distances are estimated by the RSSI values using curve fitting. We calculated the mode, average, median and maximum values of RSSI and suitable curves are chosen for them for producing best accuracy. In the WSNs, the ranges of technologies (as the LTE, Wi-Fi, ZigBee, and Bluetooth have the coverage up to 30 km, 100m, 100m and 10m, respectively [11]) and the choice of RSSIs (i.e., mode, average, median, and maximum, etc.) are the leading causes of different types of path loss. However, various types of data cannot be represented by a common model [12], and it does not explain each issue adequately. A mathematical expression can represent the fundamental shape concerning the connection between two variables at a glance. Because of the high dependency between RSSIs and distances, the relationships also cannot be represented by a linear function. Thus, the nonlinear curve fitting (NCF) is a compelling alternative as it provides more flexibility for choosing a best-fitted function from an extensive range of functions. Such NCF is based on an iterative algorithm, which varies the parameters significantly and re-evaluates until it finds the best fit. The algorithm of NCF makes it notably better than the least square and smoothing technique [13]. This curve fitting-based algorithm is so appropriate because it can represent the nature of data within a moment. For establishing the type of variables, different curves are fitted to various types of RSSIs in MATLAB. Among them, the most probable (i.e., mode) RSSIs are chosen because mode RSSI provides significantly better positioning accuracy than average, median, and maximum values. In the experiment, the mode RSSI was more than 40% for every measurement. It is known to us that signal strength is inversely proportional to distance. Therefore, only the most frequent RSSIs to each position more accurately validated the efficiency of measurements. In many cases, the rest of 60% RSSIs changed the average and median values, because from time to time, they have gone up and down to the mode. For each position, maximum RSSI has maintained a massive difference with others, and often the same maximum values have been seen for the two or more consecutive distances. The significant results on curves exhibit that the gaussian function (GAF) has most perfectly fitted to mode values and raised the appropriateness of measurement. Because of the GAF has benefits that, if data does not have a gaussian but can be transformed to obtain it fit a Gaussian. Thus, most data scientists recommended it. In the indoor experiment, ten Wi-Fi routers were kept on a floor of an office building at different locations, with a known distance from each other. Furthermore, from separate random positions, a large amount of RSSI samples was collected to the smartphone from all routers simultaneously by a Wi-Fi analyzer APP. The RSSIs to the smartphone from different routers indicate the original distances. Besides it, to verify the nature of data, an outdoor experiment was conducted in an open location with the same instruments and manners. First, we draw the curves for different RSSI values in the indoor environment. Then we

have compared the positions with the outdoor values to see the error levels. In this work, our intention to achieve the best accuracy using the curve fitting to different RSSI values. The experiment results are almost similar for both indoor and outdoor environments [14, 15]. Different results demonstrate that the proposed algorithm is more accurate, useful, reliable, and proper than other methods. The rest of this paper is summarized as follows. In Sect. 2, we have drawn the system model. Then the performance assessment of indoor localization using smartphone is discussed in Sect. 3. In Sect. 4, the experimental results are presented. Finally, in Sect. 5, we have concluded our work.

2 System Model

Consider a set of N IEEE 802.11 Wi-Fi routers, $AP = \{AP_1, AP_2, \dots, AP_N\}$ with known locations $\theta_i = [x_i, y_i]$, ($\theta_i \in \mathbb{R}^2$) for $i = 1, 2, \dots, N$, and from different random positions, $\theta_u = [x_u, y_u]$, the smartphone received RSSIs, for $u = 1, 2, \dots, M$. The positions of the smartphone situated inside the intersection of coverage of every AP. Their original distances are calculated by, $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$, ($\forall i$ and j), where d_{ui} is the Euclidean distance between the u th position and the i th location. From any position θ_u , the distances of APs are $d_{u1}, d_{u2}, \dots, d_{uN}$, with some of them are equal length. If r RSSI samples are collected to the position θ_u , with distance d_{ui} , then the set of various RSSIs of i th AP is written as $\text{RSSI}_{ui}(k) = \{\text{RSSI}_{ui}(1), \text{RSSI}_{ui}(2), \dots, \text{RSSI}_{ui}(r)\}$ for $k = 1, 2, \dots, r$. These r number of samples give x ($x \in \mathbb{N}$) types of RSSIs, where $\text{RSSI}_{ui}(1)$ occurred a times, $\text{RSSI}_{ui}(2)$ occurred b times, ..., $\text{RSSI}_{ui}(R)$ occurred q times, $a + b + \dots + q = r$. Information on the occurrence of RSSIs provided by the set $S = \{a, b, \dots, q\}$. If $\max\{S\} = l$, ($l \in S$) then the probability of most frequently occurred RSSI of i th AP is $P(S_{max}) = \frac{l}{r}$. By following the same process, the total N probabilities are taken from each position θ_u . Due to uncertainty, the occurrences of various RSSIs are changed from time to time. So, the values of a, b, \dots, q are also changed, and their numbers may be increased or decreased in different circumstances. Thus for different M positions θ_u , total $M \times N$ number of most probable RSSIs are collected along with original distances. These probabilities formed a $M \times N$ matrix, which brings to the easiest computational process. The matrix $P_{\max}^{\text{RSSI}}(d_{ui})$ describes the probabilities of mode RSSIs with original distances d_{ui} , for $u = 1, 2, \dots, M$ and $i = 1, 2, \dots, N$.

$$P_{\max}^{\text{RSSI}}(d_{ui}) = \begin{bmatrix} P(d_{11}) & P(d_{12}) & \dots & P(d_{1N}) \\ P(d_{21}) & P(d_{22}) & \dots & P(d_{2N}) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ P(d_{M1}) & P(d_{M2}) & \dots & P(d_{MN}) \end{bmatrix}$$

The similar procedures have applied for average, median, and maximum RSSI values. With the appropriate values from the above matrix, a suitable two-dimensional function $d = \psi(\text{RSSI})$ has created for each type of RSSI. Best of knowledge, the GAF (1), and these following exponential functions, EXF1 (2), EXF2 (3), and EXF3 (4), are most accurately fitted to the mode, average, median, and maximum RSSIs, respectively, since GAF produced similar results of EXF, for average, median, and maximum values. Therefore, for comparison, different exponential functions are chosen with the highest accuracy [16]. The existing signal intensity attenuation model (SIAM) (5) is also taken into consideration. Based on observation and analysis, it can say that the Gaussian estimation gives the most reliable efficiency than other mapping and model yield methods [17].

The functions, which have drawn earlier by neglecting the catalyst factors, are two-dimensional. However, if the functions are thought along with several catalyst factors, they can be written as $\phi(\text{RSSI}, c_1, c_2, \dots, c_n)$, where c_1, c_2, \dots, c_n are the catalyst variables from different objects of the environment. It is impossible to know about the expression of this function because the information of these catalyst factors is not available. It also not known to us that the catalyst factors are finite or infinite ($n \rightarrow \infty$). If the STA scans the distinct r RSSI samples of a single AP [18], with the time intervals of different counts, then for these subsequent counts, also a variety of RSSIs is seen. This scanned RSSI_{ij} of i th count and at t_j second illustrated in Table 1. For any of this count, the 1st RSSI is scanned by the smartphone at t_1 second, and the consecutive next j th RSSIs are scanned at $t_j (= t_1 + (j - 1)\delta)$ second, for $j = 1, 2, \dots, r$. For any count, the interval between two scans is denoted by δ , where $2 \leq \delta \leq 4$, ($\delta \in \mathbb{N}$). The change in RSSIs is also observed for the 1st, 2nd, ..., p th counts at any t_j second because the catalyst factors vary from time to time. For these above reasons, the expression of the catalyst function is changed at different times, which harm the measurements [19–21].

Table 1 RSSIs from an AP to a STA according to the time of scan for different counts

Time (s)	t_1	$t_2 (= t_1 + 1\delta)$...	$t_r (= t_1 + (r - 1)\delta)$
1st Count	RSSI_{11}	RSSI_{12}	...	RSSI_{1r}
2nd Count	RSSI_{21}	RSSI_{22}	...	RSSI_{2r}
.
.
.
p th Count	RSSI_{p1}	RSSI_{p2}	...	RSSI_{pr}

3 Performance Assessment

For the evaluation of the proposed algorithm, the indoor experiment has been conducted on the third floor of a four-storied office building. Ten $6\text{ m} \times 6\text{ m}$ rooms ($R1$ to $R10$) and a $30\text{ m} \times 3\text{ m}$ corridor (C) are taken as the working region, where the total area was $30\text{ m} \times 15\text{ m}$.

Rooms and the corridor have been separated by concrete walls, but each room has a wooden door and windows. Throughout this experiment, most of the inner portion of each room was filled with iron cupboard, electronic gadgets, wooden tables, etc. On the other hand, the corridor portion was free from these instruments. To conduct the work, a total of ten Wi-Fi APs was distributed among these rooms and the corridor with known locations, and the smartphone collected the RSSIs simultaneously. For the calculation, a total of 100 RSSI samples was taken from each AP. Signals of APs were propagated through both line-of-sight (LOS) and non-line-of-sight (NLOS) medium, and during experiment temperature of this floor was around 34°C .

The APs (AP 1 to AP 10), which are drawn in Fig. 1, were connected to power supply for generating the radio signals. The smartphone was kept in consecutive twelve random positions (1st to 12th) to know about the signal strength of individual APs. The APs operated on 2.4 GHz frequency band with maintaining the IEEE 802.11b/g/n standard and their communication ranges are up to 32 m. Each AP has different frequency channels and speeds. Receiver sensitivity of RSSIs is measured in decibels relative to one milliwatt (dBm), and at the time of these experiments, minimum -98 dBm sensitivity and a maximum of 20 dB variance were observed. The distances were measured with Euclidean geometry, as well as the measuring tape. Including the ten APs locations and twelve positions of STA, a total of $10 \times 12 = 120$ actual distances was measured with corresponding various types of RSSIs. No significant changes were seen in the RSSIs if the smartphone is shaken within a small interval, continuously. The experimental data have been collected at the time of human activities.

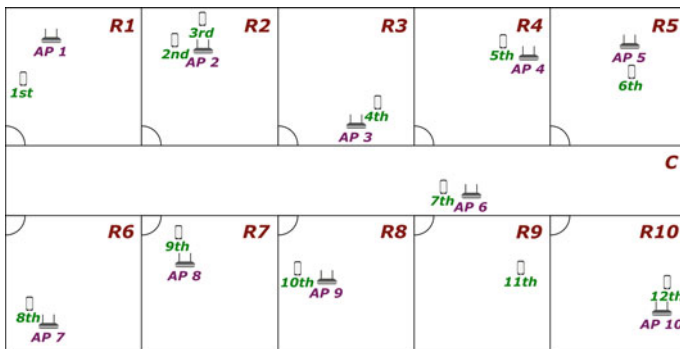


Fig. 1 Experiment scenario of the indoor region with ten rooms and a corridor, where the quarter of a circle represents the door of each room

Table 2 Coefficients of GAF, EXF1, EXF2, EXF3, and SIAM functions with coefficient of determination (R^2), sum of squared error (SSE), and root mean square error (RMSE)

GAF	EXF1	EXF2	EXF3	SIAM
$a = 2415$	$f = 0.004733$	$k = -0.03857$	$r = 0.01853$	$RSSI_d = -98$
$b = -210.4$	$g = -0.09071$	$l = 2.253$	$s = 0.9213$	$\gamma = 2$
$c = 38.247$	–	–	–	$d = 31.2046$
$R^2 = 0.9993$	$R^2 = 0.9963$	$R^2 = 0.9929$	$R^2 = 0.9766$	$R^2 = 0.3028$
$SSE = 3.725 \text{ m}^2$	$SSE = 18.56 \text{ m}^2$	$SSE = 35.54 \text{ m}^2$	$SSE = 118.2 \text{ m}^2$	$SSE = 3520 \text{ m}^2$
$RMSE = 0.198 \text{ m}$	$RMSE = 0.4397 \text{ m}$	$RMSE = 0.6085 \text{ m}$	$RMSE = 1.109 \text{ m}$	$RMSE = 5.993 \text{ m}$

The coefficients, coefficient of determinations, SSEs, and RMSEs of GAF, EXF1, EXF2, EXF3, and SIAM, respectively, are denoted in Table 2. The coefficient of determination (R^2) assesses how well a curve explains and predicted by the scattered data points around the curves. The R^2 more than 0.97 indicates that 97% of the variation in distances explained by the RSSIs, which means every curve (except, SIAM) is well fitted to these variables. The SSEs and RMSEs tell us how much the data concentrated around the curve of best fit. The RMSE gives relatively high weight than SSE for enormous errors. Based on the comparison, it concludes that the use of a curve fitting technique (CFT) with GAF is most beneficial, because the characteristic of GAF is ubiquitous in natural phenomena, as well as the parametric methods of GAF are robust and well understood. In (1), a is the height of the peak of the curve, b is the peak’s center, c is the standard deviation, and $RSSI_m$ is the measured RSSI values. In (2), g is the multiplier of measured RSSI and f is the positive real number not equal to 1. The curves (3) and (4) are similar to (2). In (5), d is the initial distance and $RSSI_d$ is the initial RSSI value at the distance d .

$$d_{\text{mod}} = a \times e^{\frac{-(RSSI_m - b)^2}{2c^2}} \tag{1}$$

$$d_{\text{avg}} = f \times e^{(g \times RSSI_m)} \tag{2}$$

$$d_{\text{med}} = 10^{(k \times RSSI_m - l)} \tag{3}$$

$$d_{\text{max}} = r \times s^{RSSI_m} \tag{4}$$

$$d_{\text{siam}} = d \times 10^{\frac{(RSSI_d - RSSI_m)}{10^\gamma}} \tag{5}$$

$$RSSI = 10 \times \log_{10} \left[\frac{P_r}{1 \text{ mW}} \right] \tag{6}$$

$$P_r = u \times d_m^v \tag{7}$$

By applying the CFT to SIAM [22], it observed that the path loss exponents of mode, average, median, and maximum RSSIs are 0.3616, 0.3825, 0.3812, and 0.4703, respectively. These path loss exponents are the same for any choice of initial distance with corresponding RSSIs. However, in the free space, path loss exponent (γ) is close to 2 [23, 24]. These two different characters of information are contradictory and cannot be possible simultaneously. Therefore, the CFT is not suitable for the SIAM with initial values, and it gives less accuracy than the mentioned GAF and EXFs. The SIAM provides its best results for mode values if γ is selected 2, along with the terminal distance and RSSI, but accuracy is next to nothing. The shape of different functions of various types of RSSIs can be observed in Fig. 2. The average and median RSSIs are relatively close to mode than maximum values. Maximum RSSIs give enormous errors and, for large distances, always have kept a minimum of 10 dB variance with the mode values. However, the variances among the mode, average, and median are not too far. The slope of the tangent of each function represents how fast the signals attenuated. Furthermore, slopes are calculated by the derivatives of these functions. Notably, the inverse, $RSSI = \psi^{-1}(d)$ of proposed GAF does not provide the real values. In this experiment, the signal sensitivity was seen between 0 and -98 dBm for the distances 0–32 m. These ranges have considered as a boundary of the Wi-Fi system, although the range can be extended using a range extender antenna. In this work, these sample variables selected by neglecting the same and nearest values. Only suitable 98 values are taken from the 120 samples for curve fitting. The measured $RSSI_m$ can be standardized with mean -49.5 dBm and standard deviation 28.43 dB, and the standardized variables are $\frac{(RSSI_m + 49.5)}{28.43}$, for $RSSI_m = -1, -2, \dots, -98$. For the calculation, only absolute errors are taken.

It is clear that the RSSI has relationship with distance as well as receive power, where P_r is receive power in milliwatt (mW). That means, distance and receive power must have a relationship. It is observed that the power function most accu-

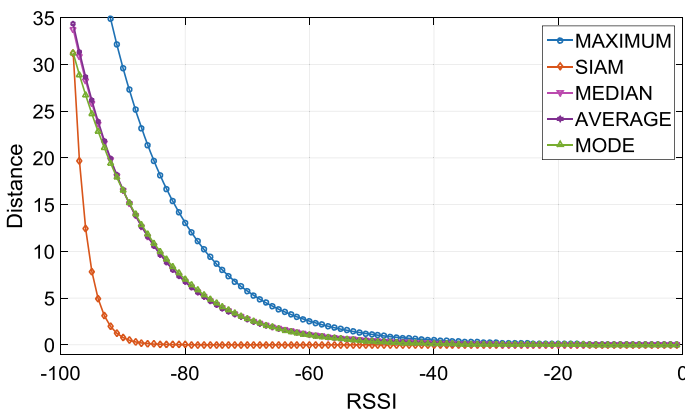


Fig. 2 Graphical representation of different functions for mode, average, median, maximum RSSIs, and SIAM in the indoor

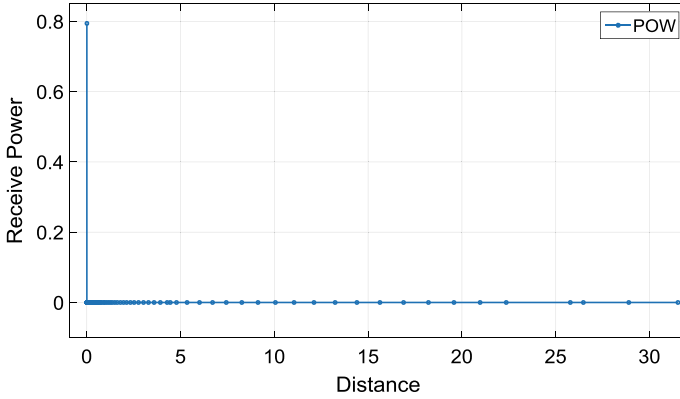


Fig. 3 Graphical relationship between receive power and distance by the power function

rately estimates the received power by distance, which is significantly better than other functions. The solution of two different types of the nonlinear equation is too complicated. For example, if Eqs. (1) and (6) are solved for receive power and distance, a complex equation will be created. So, without any complexity, CFT gives an accurate way of generating a suitable equation for any estimation. Figure 3 represents the graphical relationships of power (POW) function (7), where $u = 1.494 \times 10^{-6}$, $v = -2.652$ and d_m is the measured distance.

4 Experimental Results

For the performance verification of the proposed CFT scheme, the outdoor experiment has been conducted on a roadside ground. Multifloor buildings have surrounded the ground, and vehicles were moving through the road at the time of the experiment. The different composition has been made for the APs locations, and smartphone positions contrast to the indoor, but the total area remained equivalent, Fig. 4.

Following the similar process of indoor, a total of 98 values has been taken for the calculation. Then the original distances of outdoor have been compared with the estimated distances of indoor, which are represented by different functions. The compared results show that the SSEs and RMSEs of indoor and outdoor are similar. It is notable that, each of these experiments has done only using Wi-Fi technology.

For GAF, EXF1, EXF2, EXF3, and SIAM, the maximum errors are observed, respectively, for indoor 1.0871 m, 2.8271 m, 2.2605 m, 3.22 m, and 3.02 m and, outdoor 1.1271 m, 2.6271 m, 2.0605 m, 3.22 m, and 3.03 m. Because of enormous errors for EXF3 and SIAM, up to 3.22 m, errors are considered for the calculations. Errors difference of GAF with EXF1 and EXF2 are almost 1 m. Though the GAF and EXF have similar capabilities to fit with mode, average, median, and maximum RSSIs but each of the EXF does not fit well with other types of RSSIs. Because of having a

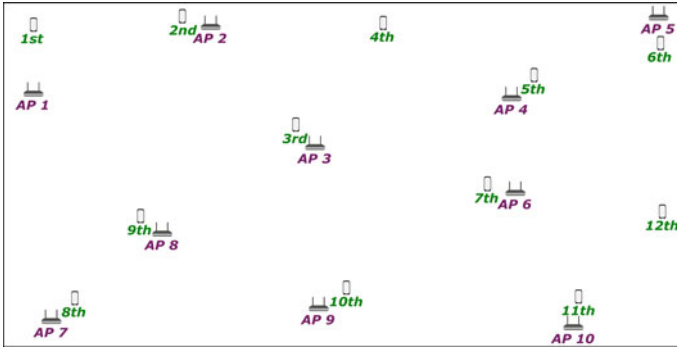


Fig. 4 Experiment scenario of the outdoor region

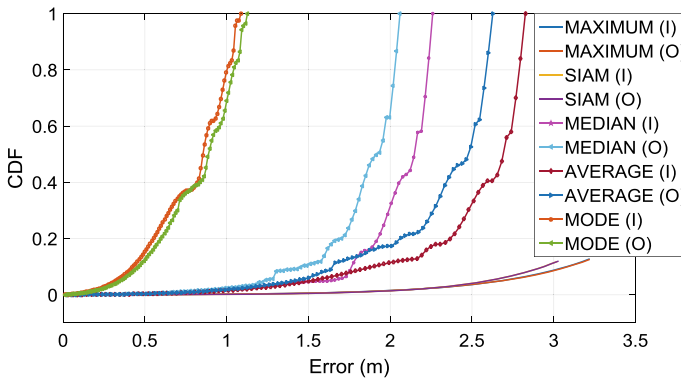


Fig. 5 Comparison of location accuracy of five curves in terms of CDF for indoor (I) and outdoor (O)

square term to the power of GAF, it provides excellent results than the others. The accuracy of location error is measured by cumulative distribution function (CDF), which is the difference between actual and estimated locations, and arranged in increasing order. From the results of Fig. 5, it is recognized that the proposed algorithm achieves significantly better positioning accuracy compared to other mapping and model-based algorithms. The CDFs of errors describe that the GAF with mode RSSIs gave the most crucial accuracy than others and almost remain similar for both indoor and outdoor environments. Different types of experiment results ensure that the mode RSSIs are most significant for position localization [25, 26], and GAF finishes this job with perfect accuracy. Besides, the better transmit power validates the excellent RSSI quality for producing the most desirable positioning accuracy.

5 Conclusions

This paper presents a Wi-Fi AP localization [27, 28], algorithm that is simple, robust, energy-efficient, affordable as well as applicable in anonymous environments [29, 30]. The proposed algorithm based on the analysis of characteristics of RSSIs and the use of proper curve fitting. Remarkably the Gaussian function provides better results than other exponential functions and can fit with any sample. Experiment results confirm that the positioning accuracy of the mentioned algorithm is notably better for measurements and it could be an efficient solution for indoor and outdoor applications [31, 32].

Acknowledgements The authors are very much thankful to Jadavpur University for giving them proper environments to make this work successful. The authors also like to thank UGC (University Grant Commission) and CSIR (Council of Scientific & Industrial Research), India for providing necessary supports.

References

1. H. Ahn, W. Yu, Environmental-adaptive RSSI-based indoor localization. *IEEE Trans. Autom. Sci. Eng.* **6**(4), 626–633 (2009)
2. B.N. Hood, P. Barooah, Estimating DoA from radio-frequency RSSI measurements using an actuated reflector. *IEEE Sens. J.* **11**(2), 413–417 (2011)
3. M. Malajner, P. Planinsic, D. Gleich, Angle of arrival estimation using RSSI and omnidirectional rotatable antennas. *IEEE Sens. J.* **12**(6), 1950–1957 (2012)
4. A.U. Ahmed, M.T. Islam, M. Ismail, Estimating DoA from radio frequency RSSI measurements using multi-element femtocell configuration. *IEEE Sens. J.* **15**(4), 2087–2092 (2015)
5. X. Ding, Y. Tian, Y. Yu, A Real-Time Big data gathering algorithm based on indoor wireless sensor networks for risk analysis of industrial operations. *IEEE Trans. Ind. Inf.* **12**(3), 1232–1242 (2016)
6. F. Martinelli, A robot localization system combining RSSI and phase shift in UHF-RFID signals. *IEEE Trans. Cont. Syst. Technol.* **23**(5), 1782–1796 (2015)
7. X. He, A. Pandharipande, Location-based illumination control access in wireless lighting systems. *IEEE Sens. J.* **15**(10), 5954–5961 (2015)
8. S.Y. Cho, Measurement error observer-based IMM filtering for mobile node localization using WLAN RSSI measurement. *IEEE Sens. J.* **16**(8), 2489–2499 (2016)
9. W. Xue, W. Qiu, X. Hua, K. Yu, Improved Wi-Fi RSSI Measurement for Indoor Localization. *IEEE Sens. J.* **17**(7), 2224–2230 (2017)
10. A.S. Paul, E.A. Wan, RSSI-based indoor localization and tracking using sigma-point Kalman smoothers. *IEEE J. Select. Top. Sig. Proc.* **3**(5), 860–873 (2009)
11. Y. Mehmood, F. Ahmad, I. Yaqoob, A. Adnane, M. Imran, S. Guizani, Internet-of-things-based smart cities: recent advances and challenges. *IEEE Commun. Magaz.* **55**(9), 16–24 (2017)
12. P.K. Sahu, E.H. Wu, J. Sahoo, DuRT: Dual RSSI trend based localization for wireless sensor networks. *IEEE Sens. J.* **13**(8), 3115–3123 (2013)
13. S. Mahfouz, F. Mourad-Chehade, P. Honeine, J. Farah, H. Snoussi, Non-parametric and semi-parametric RSSI/distance modeling for target tracking in wireless sensor networks. *IEEE Sens. J.* **16**(7), 2115–2126 (2016)
14. A. Maddumabandara, H. Leung, M. Liu, Experimental evaluation of indoor localization using wireless sensor networks. *IEEE Sens. J.* **15**(9), 5228–5237 (2015)

15. S.K. Gharghan, R. Nordin, M. Ismail, J.A. Ali, Accurate wireless sensor localization technique based on hybrid PSO-ANN algorithm for indoor and outdoor track cycling. *IEEE Sens. J.* **16**(2), 529–541 (2016)
16. R.M. Sandoval, A. Garcia-Sanchez, J. Garcia-Haro, Improving RSSI-based path-loss models accuracy for critical infrastructures: a smart grid substation case-study. *IEEE Trans. Ind. Inform.* **14**(5), 2230–2240 (2018)
17. Q. Luo, Y. Peng, J. Li, X. Peng, RSSI-based localization through uncertain data mapping for wireless sensor networks. *IEEE Sens. J.* **16**(9), 3155–3162 (2016)
18. S. Mazuelas et al., Robust indoor positioning provided by real-time RSSI values in unmodified WLAN networks. *IEEE J. Select. Top. Sig. Process.* **3**(5), 821–831 (2009)
19. C. Pu, W. Chung, Mitigation of multipath fading effects to improve indoor RSSI performance. *IEEE Sens. J.* **8**(11), 1884–1886 (2008)
20. L. Lin Shen, W. Wong Shung Hui, Improved pedestrian dead-reckoning-based indoor positioning by RSSI-based heading correction. *IEEE Sens. J.* **16**(21), 7762–7773 (2016)
21. N. Jain, S. Verma, M. Kumar, Adaptive locally linear embedding for node localization in sensor networks. *IEEE Sens. J.* **17**(9), 2949–2956 (2017)
22. H. Kwasmeh, S. Ekin, RSSI-based localization using LoRaWAN technology. *IEEE Access* **7**, 99856–99866 (2019)
23. Y. Gu, F. Ren, Energy-efficient indoor localization of smart hand-held devices using bluetooth. *IEEE Access* **3**, 1450–1461 (2015)
24. C.L. Nguyen, U. Raza, LEMOn: wireless localization for IoT employing a location-Unaware mobile unit. *IEEE Access* **7**, 40488–40502 (2019)
25. J. Yoo, Change detection of RSSI fingerprint pattern for indoor positioning system. *IEEE Sens. J.* **20**(5), 2608–2615 (2020)
26. P. Savazzi, E. Goldoni, A. Vizziello, L. Favalli, P. Gamba, A wiener-based RSSI localization algorithm exploiting modulation diversity in LoRa networks. *IEEE Sens. J.* **19**(24), 12381–12388 (2019)
27. V. Bianchi, P. Ciampolini, I. De Munari, RSSI-based indoor localization and identification for ZigBee wireless sensor networks in smart homes. *IEEE Trans. Instrum. Meas.* **68**(2), 566–575 (2019)
28. J. Koo, H. Cha, Localizing WiFi access points using signal strength. *IEEE Commun. Lett.* **15**(2), 187–189 (2011)
29. Y. Yao et al., Multi-channel based sybil attack detection in vehicular Ad Hoc networks using RSSI. *IEEE Trans. Mob. Comput.* **18**(2), 362–375 (2019)
30. D. Konings, F. Alam, F. Noble, E.M. Lai, Device-free localization systems utilizing wireless RSSI: a comparative practical investigation. *IEEE Sens. J.* **19**(7), 2747–2757 (2019)
31. J. Yan, H. Zhao, X. Luo, C. Chen, X. Guan, RSSI-based heading control for robust long-range aerial communication in UAV networks. *IEEE Internet Things J.* **6**(2), 1675–1689 (2019)
32. A. Booranawong, N. Jindapetch, H. Saito, Adaptive filtering methods for RSSI signals in a device-free human detection and tracking system. *IEEE Syst. J.* **13**(3), 2998–3009 (2019)

A Novel Miniaturized Fractal Antenna on Circular Patch with Defected Ground Structure (DGS)



Sakshi Singh, Jitesh Kumar, Arvind Gaur, and Rahul Vishnoi

Abstract A novel miniaturized fractal antenna on circular patch with DGS is examined in this paper. The designed antenna will resonate in X band, Ku band, and K band at 8.91 GHz, 14.27 GHz, 18.58 GHz, and 22.31 GHz frequencies. A dumbbell shape DGS structure is introduced to enhance bandwidth, and the newly improved antenna will have the application in Ka band also. The improved antenna will resonate at 8.85 GHz, 14.19 GHz, 18.65 GHz, 22.43 GHz, and 26.89 GHz. The designed antenna is miniaturized and having the substrate area of $14.3 \text{ mm} \times 14.3 \text{ mm}$ with a circular shape patch of radius 4.2 mm. FR4 (4.4) material is used to design the substrate of antenna. The examined miniaturized antenna is compatible with satellite communication, radio location, mobile communication, space research, and radio navigation. The simulation of investigated antenna is performed on Ansoft HFSS antenna design simulation software.

Keywords Fractal antenna · Defected ground structure (DGS) · Circular patch

1 Introduction

The advance communication system requires compact multiband antenna which can produce high gain and wide bandwidth. Circular patch is used in proposed design to miniaturize the antenna. The substrate is square in shape and made up of FR4 material. The examined antenna is very compact and has the total area of $14.3 \text{ mm} \times 14.3 \text{ mm}$. Several techniques have been used to miniaturize the antenna like different types of patches are used to achieve compact and reliable antenna. In modern days, antenna operating at multiple frequencies are desirable; it also saves the device space and hence manufacturing cost of the system. To get rid of this problem, we use a single antenna operates at multiple frequencies [1]. Fractal geometries are used to form

S. Singh (✉) · R. Vishnoi
Faculty of Engineering, TMU, Moradabad, India

J. Kumar · A. Gaur
University Polytechnic, TMU, Moradabad, India

multiband antennas, fractal technique is used because it has self-similar property, and circular patch is used to reduce the effective area of antenna. Square slots have been cut in the form of fractal geometry in iterative fashion to achieve multiband [2]. Different fractal geometries have been used for different applications. Other antenna with circular slots on circular patch has been designed in fractal fashion, and antenna can be used for satellite communication. The antenna reported in [3] having the fractal geometry similar to the proposed one, but it is not very much compact and do not have any improved version. Many techniques have been used to minimize the dimensions of antenna like a dielectric semicircle arc which is projected on the antenna [4] and is one of them. The patch size is made smaller by enhancing the arc and by increasing the projection. Size reduction is done without any degradation in the antenna gain. Different types of fractal geometries have been introduced to achieve multiband and high gain. Smith chart geometry design slot on circular patch is introduced in [5].

Several more antennas with circular patch have been proposed in literature. Authors in [6] proposed antenna to reduce cross polarization, and the antenna is a type of probe fed to patch microstrip patch antenna. Antenna with circular substrate have been presented and fabricated in [9, 10]. Annular ring slot is designed on circular patch because of ease of fabrication [7, 8], and monopole radiation pattern is achieved in [7, 9]. Circular patch antenna with circular polarization for 5G application is designed by the author in [10]. Defected ground structure (DGS) geometries on circular patch is also investigated in [11], in which smallest DGS structure is used to reduce the cross-polarized radiation. DGS are used in numerous passive and active devices. Every DGS shape has different impact on the antenna working, and it depends on the size and geometry of DGS structure. DGS has its own applications like miniaturizing component size, suppressing cross-polarization, and higher-order modes [12]. Here, dumbbell shape DGS structure is used to enhance bandwidth of fractal antenna, and using DGS structure shows the performance of antenna in Ka band also. One research in [13] is done with hexadecagon circular patch to form a dual-band antenna. The proposed antenna gives good return loss and high gain which is achieved by using defective ground structure. This antenna can be used for satellite communication in Ku band. In [14], authors have designed different shape DGS antennas like pentagonal shape patch antenna having pentagonal DGS, pentagonal shape, circular shape type patch antennas having square DGS, pentagonal DGS, and circular shape DGS which have been investigated. This antenna can be utilized in wireless communication systems. The author in [15] uses H-shape DGS to improve the bandwidth, and antenna is formulated for Wi-Fi applications. Relative study of rectangular patch and circular shape patch antennas have been studied and investigated that the circular patch antennas has higher bandwidth and less side lobe power as compare to rectangular patch antenna [16]. In [17], author investigated probe-fed type circular shape patch antenna having ground plane, and then, in second investigation, dot shape DGS is used in minimizing cross-polarization level of electric field, and this was studied experimentally. Different shapes of DGS like E shape, F shape, and U shape have been studied and compare. The investigated antenna gives tri-band and enhances the gain of antenna [18]. In [19], author gives a new antenna which

consists of concentric circular rings, and metallic type of shielding is done on the back of antenna to eliminate any radiation or any leakage. The proposed designed antenna is simulated, and results are measured [20–22]. The antenna gives wide stop band characteristics.

2 Specifications for Antenna Designing

Circular type patch antenna is broadly used antennas in other microstrip antennas. It can resonates at multiband frequency ranges. The radius of used circular patch can be moderate and can be calculated using Eq. (1)

$$R = \frac{F}{\left(1 + \frac{2h}{\pi \epsilon_r F} \left(\ln \frac{\pi F}{2h}\right) + 1.7726\right)^{1/2}} \tag{1}$$

where

$$F = \frac{8.791 \times 10^9}{\epsilon_r f_r}$$

Effective radius of circular patch is

$$R_e = R \left(1 + \frac{2h}{\pi \epsilon_r R} \left(\ln \frac{\pi R}{2h}\right) + 1.7726\right) \tag{2}$$

And resonance frequency is given by

$$f_r = \frac{1.8412C_0}{2\pi R_e \sqrt{\epsilon_r}} \tag{3}$$

The ground plane is taken as square of size 14.3 mm × 14.3 mm and FR4 (4.4) is used as the substrate material. The proposed antenna is a type of fractal antenna for which circular patch is used for designing. The height of the substrate is 2 mm, and the radius of used circular patch is taken as 4.2 mm. Microstrip type of feed line method is used to feed antenna. The simulation results shows that the antenna can resonates in X band, Ku band, and K band. Further, a dumbbell shape DGS structure is implemented in ground to improve the bandwidth of antenna, and the improved antenna will resonates in Ka band also as shown in Table 1.

Table 1 Antenna parameters with description

Parameter	Description	Value (in mm)
L_g	Ground plane length	14.3
W_g	Ground plane width	14.3
L_s	Length of substrate	14.3
W_s	Width of substrate	14.3
L_f	Feed line length	2.95
W_f	Feed line width	0.9
h	Substrate height	2
R	Patch radius	4.2

3 Antenna Design Methods

The investigated antenna is designed on circular patch of radius $a = 4.2$ mm. The ground and substrate is taken as square of size $14.3 \text{ mm} \times 14.3 \text{ mm}$. A circle of 0.84 mm is subtracted from the patch with the concept of fractal geometry in zeroth iteration. In the first iteration, four circles of radius 0.42 mm are subtracted, and in second iteration around these four circles, another 16 more circles of radius 0.21 mm are being subtracted to form final fractal geometry as shown in Fig. 1. Further, dumbbell shape DGS have been created to increase the bandwidth of investigated antenna, and this DGS structure also created one more Ka frequency band application as shown in Fig. 2.

Fig. 1 Proposed fractal antenna geometry

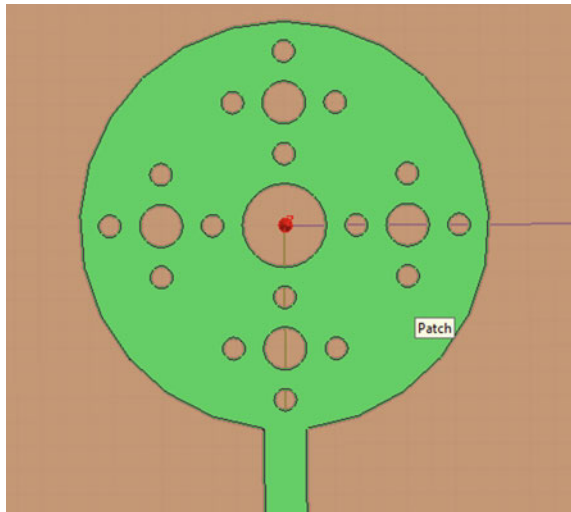
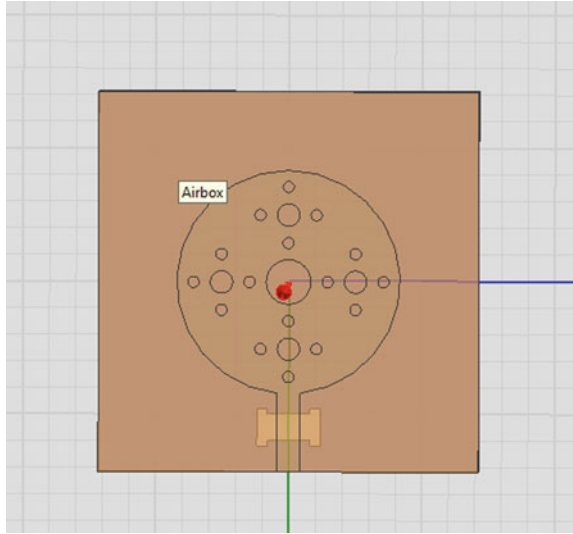


Fig. 2 Proposed fractal antenna with DGS structure



3.1 Validation of Proposed Design

Reference (GHz)	Frequency band	Material used	Dimension of antenna (mm)	Patch radius (cm)
[2]	X, Ku, K, V band	Rogers RT Duroid5880	1.92 × 1.9 2 cm ²	0.525
[3]	X, K, Ku band	Rogers RT Duroid5880	6.85 × 6.85 cm ²	1.71
[4]	S band	Plastic material	7.0 × 7.0 cm ²	1.8
[5]	L, S, C band	Rogers TMM4	7.1 × 7.1 cm ²	1.77
[6]	S band	Taconic TLP-3	6.0 × 6.0 cm ²	1.5
Proposed	X, Ku, K, Ka band	FR 4	1.43 × 1.43 cm ²	0.42

4 Results and Discussion

The investigated antenna is designed and simulated in Ansoft HFSS. This fractal antenna with DGS structure is designed to obtain X, Ku, K, and Ka band and can be used for satellite communication, radio location, mobile communication, space research, and radio navigation. Multiple iterations have been done to obtain the desired antenna. The simulation results of return loss without DGS and with DGS

are shown in Fig. 3. The zero iteration is done by creating a circular slot on circular patch. Further, four more circular slots have been cut around the center circle, this geometry is repeated to form a fractal structure, and this fractal antenna is modified by using dumbbell shape DGS which leads to add one more Ka frequency band with enhanced bandwidth.

Figure 4 shows the gain polar plot of proposed antenna structure with DGS which is the improved version of designed fractal antenna. The gain polar plot shows that antenna performance is improved in terms of gain after using DGS structure.

Figure 5 shows the gain radiation pattern of improved antenna structure. It indicates good radiation pattern for all frequency segments. Fig. 6 shows the simulated results for VSWR at measured frequencies for the improved antenna structure having DGS geometry. VSWR plot indicates good impedance matching at desired frequencies.

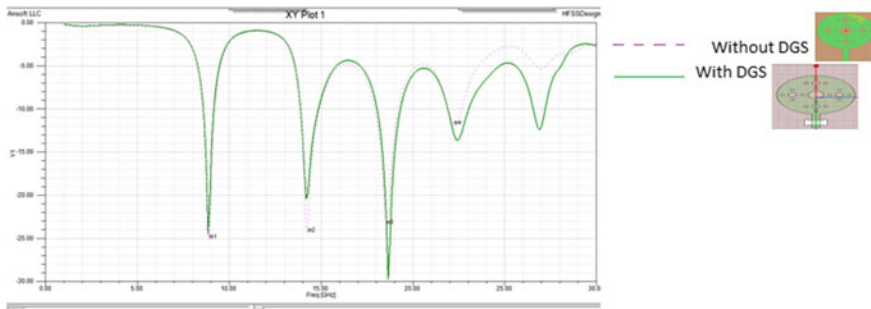


Fig. 3 Return loss of antenna design with DGS and without DGS structure

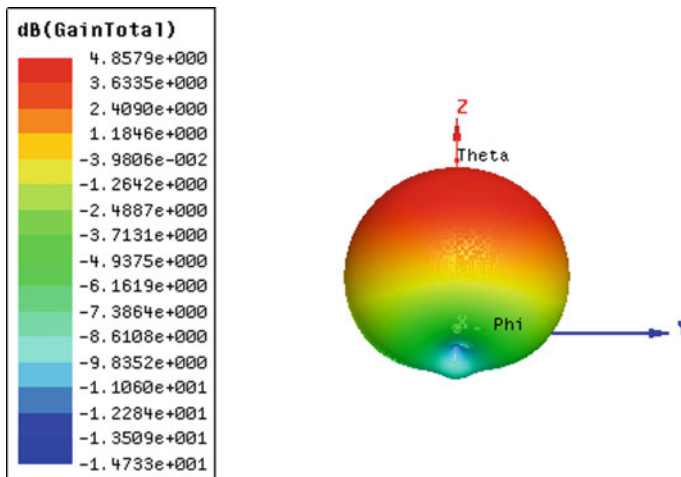


Fig. 4 Gain polar plot of antenna structure with DGS

Fig. 5 Gain radiation pattern of antenna structure

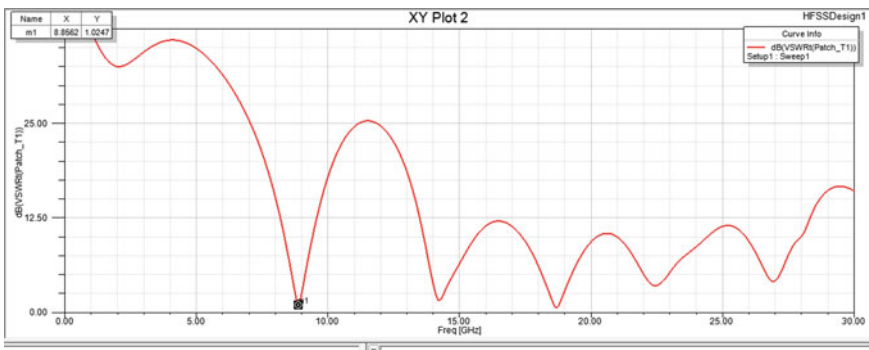
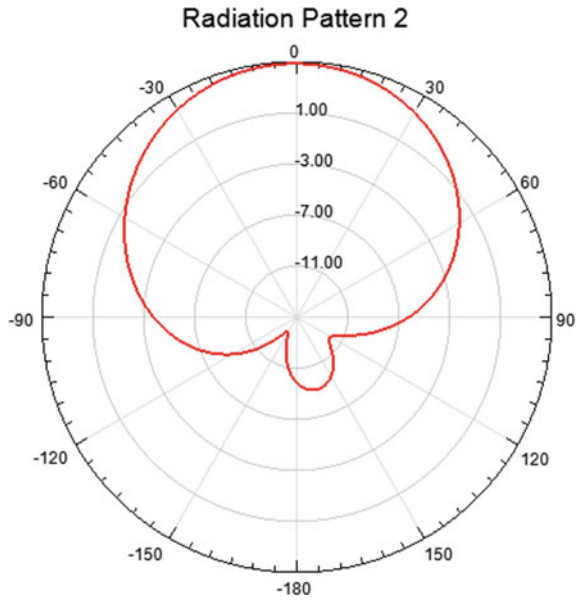


Fig. 6 VSWR plot of improved antenna with DGS geometry

The electric field and magnetic field distributions are presented in Fig. 7a, b denote the expansion of field started from center of patch. For design validation, the gain of improved version of antenna structure is shown in Fig. 8. The maximum gain of improved antenna with DGS structure is 5 dB.

The radiation pattern at different resonating frequencies of improved version of antenna structure with DGS geometry is shown in Fig. 9a–e. These figures show that antenna have different shapes of radiation pattern at different frequencies. Figure 10 shows 3D polar plot of improved antenna at resonating frequency 8.85 GHz.

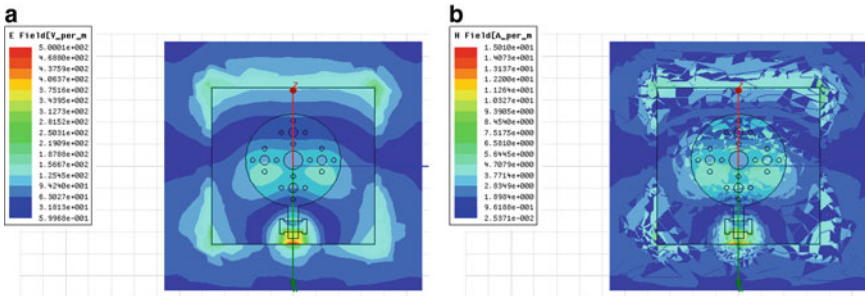


Fig. 7 **a** Electric field distribution and **b** magnetic field distribution of antenna structure with DGS geometry

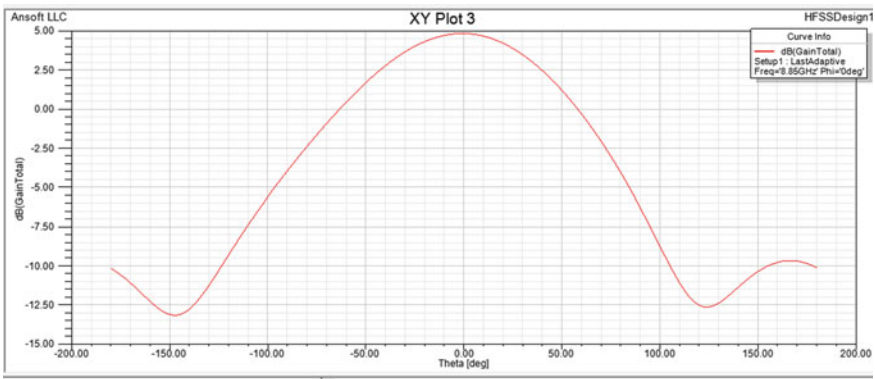


Fig. 8 Gain of improved antenna with DGS structure

5 Conclusion

The designed improved antenna structure with DGS have resonating frequency at X band, Ku band, K band, and Ka band, so it can be utilized for satellite communication, radio location, mobile communication, space research, and radio navigation applications successfully. The designed structure fulfills all the requirements for these applications and miniaturized in terms of patch size. The antenna is a type of fractal geometry; thus, it is resonating at multiple frequencies and DGS geometry added one more frequency band. The antenna has good gain characteristics which makes the proposed structure more feasible for given applications.

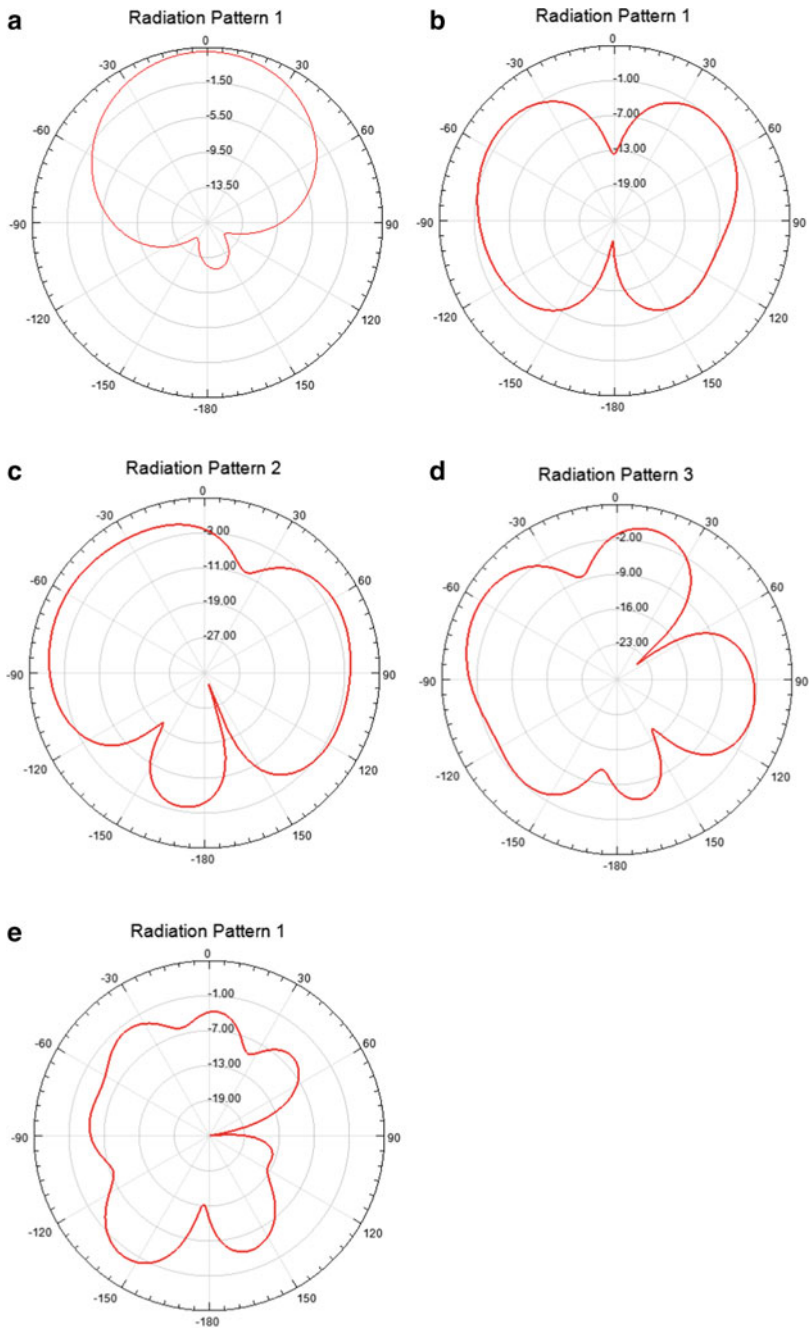


Fig. 9 a Radiation pattern at 8.85 GHz, b radiation pattern at 14.19 GHz, c radiation pattern at 18.65 GHz, d radiation pattern at 22.43 GHz, and e radiation pattern at 26.89 GHz

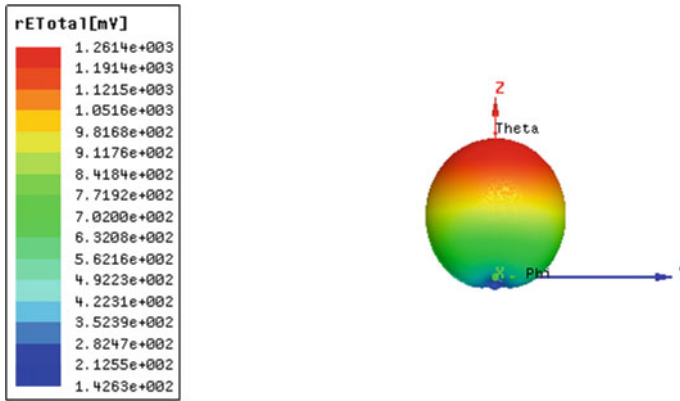


Fig. 10 3D polar plot at 8.85 GHz

References

1. C.A. Balanis, *Antenna Theory: Analysis and Design* (Wiley-Interscience, 2012)
2. B.B. Chowdhury, R. De, M. Bhowmik, A novel design for circular patch fractal antenna for multiband applications, in *3rd International Conference on Signal Processing and Integrated Networks (SPIN)* (2016)
3. D. Selvaraj, K.G. Sandeep, A.R. VeeraSudhan, N.L. Venkatesa Suhin, V. Parameshwaran, Design of circular shaped fractal antenna for the application of Satellite Communication V. Int. J. Eng. Res. Technol. (IJERT) **7**(04) (2018)
4. A. Motevasselian, W.G. Whittow, Miniaturization of circular patch microstrip antenna using an arc projection. *IEEE Antennas Wirel. Propag. Lett.*
5. N. Rao, D. Kumar, Fractal slots on circular patch inspired by smith chart for GSM, Radar, WLAN, and ITS Application V, in *2015 Loughborough Antennas & Propagation Conference (LAPC)*
6. S. Samanta, P. Soni Reddy, K. Mandal, Cross-polarization suppression in probe-fed circular patch antenna using two circular clusters of shorting pins. *Trans. Antennas Propag.* <https://doi.org/10.1109/TAP.2018.2819895>
7. A. Al-Zoubi, F. Yang, A. Kishk, A broadband center-fed circular patch-ring antenna with a monopole like radiation pattern. *IEEE Trans. Antennas Propag.* **57**(3) (2009)
8. D.M. Kokotoff, R.B. Waterhouse, C.R. Birtcher, J.T. Aberle, Annular ring coupled circular patch with enhanced performance. *Electron. Lett.* **33**(24) (1997)
9. S. Liu, W. Wu, D.-G. Fang, Wideband monopole-like radiation pattern circular patch antenna with high gain and low cross-polarization. *Trans. Antennas Propag.* <https://doi.org/10.1109/TAP.2016.2536418>
10. S. Raj, N. Kishore, G. Upadhyay, S. Tripathi, V.S. Tripathi, A Compact Design of Circularly Polarized Fractal Patch Antenna for 5G Application. 978-1-5386-8221-0/18/2018 IEEE
11. C. Kumar, D. Guha, Asymmetric and compact DGS configuration for circular patch with improved radiations. *IEEE Antennas Wirel. Propag. Lett.* <https://doi.org/10.1109/LAWP.2019.2962569>
12. M.K. Khandelwal, B.K. Kanaujia, S. Kumar, Defected ground structure: fundamentals, analysis, and applications in modern wireless trends. *Int. J. Antennas Propag.* (2017)
13. K.K. Naik, P.A. Vijaya Sri, Design of hexadecagon circular patch antenna with DGS at Ku band for satellite communications. *Progr. Electromagn. Res. M* **63**, 163–173 (2018)
14. R.K. Yadav, S. Das, R. L. Yadava, DGS based microstrip patch antennas for UWB systems, in *2013 3rd IEEE International Advance Computing Conference (IACC)*

15. S.K. Sarkar, A. Ghosh, A. Ghosh, K. Senthil Kumar, T. Datta, Implementing a power efficient PN sequence generator using spintronic devices for CDMA based wireless mobile communication. *J. Comput. Theoret. Nanosci.* **5**(6), 1–6 (2008)
16. I. Das, R.N. Shaw, S. Das, Location-based and multipath routing performance analysis for energy consumption in wireless sensor networks, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol. 661, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_59
17. A. Thatere, D. Arya, P.L. Zade, Wi-Fi microstrip antenna with DGS, in *International Conference on Innovations in Power and Advanced Computing Technologies [i-PACT2017]*
18. B. Ramesh, V. Rajya Lakshmi, Design of an Ultrawideband Circular Microstrip Antenna with DGS. 978-1-4799-3267-2/13/2013 IEEE
19. C. Kumar, D. Guha, Nature of cross-polarized radiations from probe-fed circular microstrip antennas and their suppression using different geometries of defected ground structure (DGS). *IEEE Trans. Antennas Propag.* **60**(1) (2012)
20. A. Kakkar, Nirdosh, S. Sah, A tri-band circular patch microstrip antenna with different shapes in DGS for Ku and K applications, in *2017 2nd International Conference on Telecommunication and Networks (TEL-NET 2017)*
21. D. Guha, S. Biswas, M. Biswas, J.Y. Siddiqui, Y.M.M. Antar, Concentric ring-shaped defected ground structures for microstrip applications. *IEEE Antennas Wirel. Propag. Lett.* **5** (2006)
22. D. Guha, M. Biswas, Y.M. M. Antar, Microstrip patch antenna with defected ground structure for cross polarization suppression. *IEEE Antennas Wirel. Propag. Lett.* **4** (2005)

Rotating Acoustic Reflector Parameter Trade-Off for Near-Outdoor Audio Event Detection



Ganesh Bhagwat, Sangeeth Jayaprakash, and Anirban Bhowmick

Abstract Present visual surveillance systems have limitations, like poor data in low or no lighting conditions, non-line of sight, large data storage. An audio surveillance system would better suit to overcome these limitations and complement the visual surveillance, especially for large area and outdoor conditions. In this paper, we have explored the development and use of a rotating parabolic reflector microphone for near outdoor audio recording. The low-cost rotating reflector brings out interesting trade-off between its size and its speed of rotation for both sensitivity and directionality. Specific improvement to the shape of the reflector and the type of microphone along with suitable rotational speed has shown promise of effective audio signal recording economically. It has also brought out the issue of audio event duration and its spectral content for effective detection. We demonstrate the utility of such an audio sensor system for a small-scale audio event classification task, using a hierarchical classification scheme.

Keywords Acoustic reflector · Hierarchical classification · Audio event detection · Directivity

1 Introduction

Audio event detection is a demanding task for the purpose of security and surveillance of a given premise. But when deployed in an outdoor scenario, it adds environmental and ecological uses, such as poacher detection, animal and bird call pattern study.

G. Bhagwat (✉)
Expleo India Pune, Pune, India
e-mail: ganesh.bhagwat@expleogroup.com

S. Jayaprakash
Indian Institute of Science Bangalore, Bangalore, India

A. Bhowmick
Vellore Institute of Technology, Bhopal, India
e-mail: anirban.bhowmick@vitbhopal.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_44

Although it has been acknowledged that audio event detection is harder than classification, less attention has been given to the aspect of event detection improvement of continuous audio streams. The detection of an audio event becomes more difficult when the source is far from the microphone and of unknown direction.

To improve the sensitivity and directional selectivity of the sensor, we have explored the use of an acoustic reflector along with the microphone. This does improve the signal quality and can help audio event detection of streaming audio. An acoustic reflector is a passive structure which can be used along with microphones to gather and focus sound waves from a longer distance. Many authors have also shown the use of an acoustical reflector for different applications. Yoshifumi et al. [1] have proposed a ceiling dome microphone for monitoring breath sound within a room in a quiet home. In an outdoor scenario, Aumann et al. [2] have described the use of small parabolic reflector microphone to capture the buzz of honeybees to monitor the health of beehives.

Researchers in audio event detection and classification have applied different schemes, like detection through classification framework in which an event classification model is pre-trained, and then, the test audio signal is classified by a sliding window technique [3, 4]. A hierarchical classification scheme has been used by Atrey [5] for multimedia surveillance to detect various normal and excited state human activities. Various deep learning schemes have also been exploited in Juncheng Li's [6] work. Many such works have been summarized and put together by Marco Crocco in his review paper [7].

For audio event detection, separating an event from the background is itself a challenge; Simo Moncrieff [8] has summarized different schemes to separate an event from the background; they have shown that the background model should be online and adaptive.

In this paper, we have explored a new type of sensor system of slowly rotating acoustic reflector with microphone, so that we can get a good sensitivity in all directions using only a single microphone. We show how a good combination of microphone and reflector pair can improve the event localization. We have evaluated the performance of the device for event classification around our laboratory building based on energy thresholding for detecting an event from the background and then use hierarchical classification to classify different events.

2 Audio Recording Scheme

The present experimental setup is shown in Fig. 1. We have placed the rotating sensor on our laboratory rooftop so that we can have an uninterrupted power supply and will enable monitoring of audio events surrounding our building.

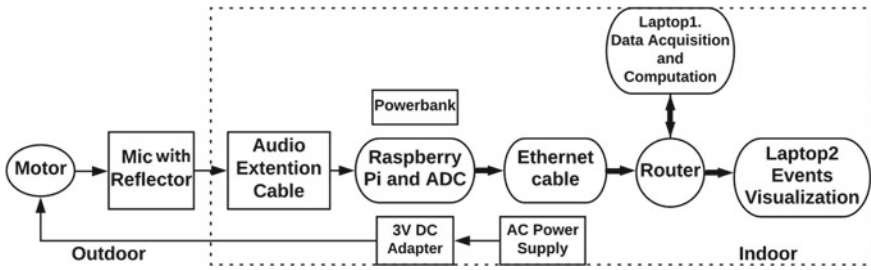


Fig. 1 Outdoor audio acquisition and indoor processing scheme

2.1 Acoustic Reflector

The inherent property of a dish microphone is similar to that of a dish antenna. The basic physical principles governing both are similar. The acoustic reflector gives an enhanced gain to the audio recording through the reflected energy. It enhances the directional gain of the microphone in the direction to which the dish is open. So after monitoring the signals for a complete rotation, we can estimate an approximate direction of a stationary sound source based on the rotation time and the angular position.

The focal position is an important parameter to get maximum amplification of the acoustic signal. The expression for focal length (a) of a parabolic dish is:

$$a = \frac{D^2}{16 * d} \tag{1}$$

where D is the dish diameter, and d is the depth of the dish.

Best material and size for the acoustic reflector had to be chosen for the experimental setup. The materials we tested are steel, plastic and aluminum. First, to ascertain the focal point of the reflectors, we have shifted the microphone position forward and backward of the theoretical focal point of each reflector. We recorded the input stimulus at 0 azimuth angle at a distance of 2 m from the reflector. The received signal power is shown in Table 2. Input stimulus is a recorded speech signal of 10 s duration and played through a loudspeaker. The experiment was carried out in a sound-treated studio. We can observe in Table 2 that the steel dish has the same theoretical and best experimental focal point, but for aluminum and plastic dish, the focal point has shifted from its theoretical focal point. The reason could be that the plastic and aluminum dishes are not exactly parabolic shape, whereas the steel dish had parabolic shape. After ascertaining the focal point of the acoustic reflectors, we performed another experiment to verify the performance of each reflector with a source placed at different distances at the 0 azimuth angle. In this experiment, the microphone was placed at the best experimental focal point, and the same earlier input stimulus is used. The experimental observation is shown in Table 1. We performed this experiment in the same sound-treated room. The dimensions and shape

Table 1 Signal power in dB according to distance and material of the dish

S.No	Dish type	Distance between audio source and reflector			
		2 m	4 m	6 m	8 m
1	Plastic (dB)	-54.16	-59.18	-61.29	-63.38
2	Aluminum (dB)	-55.19	-59.61	-61.92	-63.56
3	Steel (dB)	-55.28	-60.96	-62.21	-64.12

Table 2 Experimental focal point comparison

Dish type	Theoretical focal point (cm)	Experimented focal points (cm)	Signal power (dB)	Chosen focal point (cm)
Plastic	8.7	5	-55.43	6
		6	-54.16	
		8.7	-55.86	
Aluminum	10.1	7	-55.87	8
		8	-55.19	
		10.1	-56.42	
Steel	5	4	-57.63	5
		5	-55.28	
		6	-56.32	



Fig. 2 Acoustic reflectors, steel (dimension: D = 24.2 cm, d = 7.4 cm), plastic (dimension: D = 43.5 cm, d = 13.5 cm) and aluminum (dimension: D = 46 cm, d = 13 cm) (left to right)

of each reflector are given in Fig. 2. We can see that the dimensions of the plastic and aluminum reflectors are approximately same, and the steel reflector is slightly smaller. From Table 1, we note that the gain performance is slightly better for plastic compared to other reflectors. The steel reflector has shown good results with a smaller size. However, due to the lightweight of plastic to mount on a motor shaft and also better performance at a longer distance, we have used the plastic reflector for further processing.

Fig. 3 AED setup with mounting of reflector dish on a motor



2.2 Rotational Motor

We have mounted the acoustic reflector on a motor shaft to get a 360° rotation. The dish rotating from a motor introduces its own acoustic noise and shaft vibration. Both of these should be minimized for choosing and constructing the motor mounting. After experimentally trying different motors (DC as well as Stepper motor), DC motor with a speed upto 10 rpm has been chosen. The motor has a voltage rating of 12 V, but to reduce the rpm, we have used 3 V DC adapter. Using the 3 V adapter, we got a motor speed of 3 rpm suitable for acoustic event gathering . We have used optical encoder with the motor to get the direction of the reflector. Cloth and sponge materials are used to muffle the motor noise, and a wooden shaft with plastic rope is used to attach the reflector dish so that vibration can be minimized. The shaft was rolled in cloth, and the motor was totally enclosed in a box to reduce the radiating motor noise. This audio event detection (AED) setup at a near-outdoor environment is shown in Fig. 3.

2.3 A/D Conversion

For audio signal digitization, it is important to locate the A/D close to mic, so that electrical noise is minimized. Since the mic is rotating with the reflector, the A/D would also rotate, and hence, it should be compact, rigged and lightweight. It is also low power, well suited for open Linux software implementation. Raspberry

pi is programmed to record audio data seamlessly. Audio signal picked up by the microphone is acquired by the raspberry pi at $f_s = 16$ kHz and a chunk size of 4096. Raspberry pi sends the data to a local server to acquire data continuously from the specific port. The Raspberry pi is connected to the laptop through a long Ethernet cable. It is a point-to-point data communication scheme.

Data Logging A local area network is configured to balance the computational load and live visualization of the audio events. The server laptop-1 acquires data constantly from the Raspberry pi through a specific port and processes the data for event detection and classification. After the processing, the visualization and data logging takes place on laptop-2. The results are visualized as 3D spectrogram of the audio signal along with event labels which are refreshed every 1 min interval. In the data logger, the detected event label is logged at every 1 s interval.

3 Rotating Reflector Sensor

3.1 Directivity, Sensitivity and Motor Speed

We have performed an experiment in outdoor condition to infer the effect of motor rotation speed and source distance on the directivity. The experiment shows that if an event lasts for a full rotation of the reflector and if the event is within a distance of ≈ 100 m, then we can estimate the direction of the event with maximum confidence. But, if the distance is farther, then we cannot conclude it with sufficient confidence. We have also found that the reflector shows no directivity (omni-directional) for prayer recording, since the prayer source is at a long distance of >1 km. To assess the interaction of directivity and distance, we used a recorded speech waveform of 2 min as our distant stimulus. We have placed the reflector facing the sound source (0° azimuth) and recorded the signal till the completion of one full rotation. We have performed this experiment for different distances in the near outdoor scenario. A comparison of received signal power is shown in Fig. 4, where Δ is the sensitivity difference between maximum and minimum signal energy. We can observe the reduction in directivity as the distance increases and the difference gradually decreases. After a certain distance, the sensitivity difference is close to zero. Beyond that distance, it will be difficult to locate the direction of the source even in the stationary reflector case. Increasing the diameter of the reflector can increase the directionality, but will be infeasible as a low-cost rotational solution. We can see that reducing the motor speed is giving better directionality. As the motor speed increases, motor noise also increases, and the chance of detecting feeble events gets reduced.

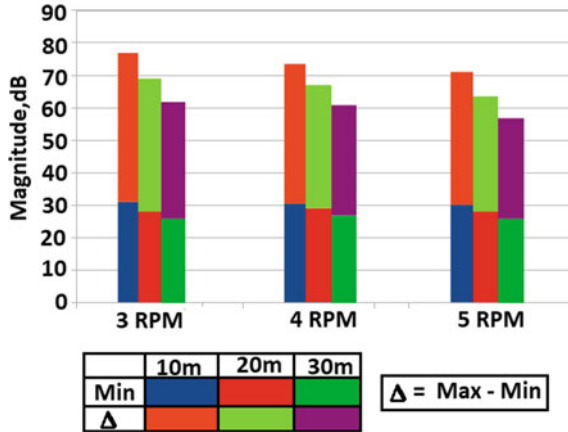


Fig. 4 Relation between motor speed, distance and directivity (outdoor). This shows that the parabolic reflector gives better gain and directivity for lower RPM

3.2 Mic and Reflector Improvement

To improve the sensitivity of the measure of difference between front and back of the reflector, we made further modification in the design. We replaced the omnidirectional microphone with a directional microphone (sub-cardioid) to see the improvement in sensitivity (we have to choose a small size, lightweight directional mic so that the rotating platform load is not affected). The mic’s directivity pattern is shown in Fig. 6b. We can observe that the difference in sensitivity (Δ) is improved in the case of unidirectional microphone. We also modified the shape of the reflector and attached a cylindrical mouth on top of the plastic parabolic reflector to give it more depth and reduce the effect of scattered sound (similar idea is used in a microwave dish antenna also). The directivity patterns of the modified reflector and the parabolic reflector are shown in Fig. 6a. We can see an improvement of 2–3 dB in the directional sensitivity because of the modified reflector. A picture of the modified parabolic reflector is shown in Fig. 5. These directionality measurements are conducted in the anechoic chamber such that the directional advantage of the recording setup is determined without the reflections from side walls. We have set up a measurement procedure with the reflector microphone placed at the center of a hypothetical circle of radius 2 m, and a loudspeaker is located around the perimeter of the circle at intervals of 30°, as well as at 0° and 180°. (exact front and back). The improved directivity of the modified parabolic reflector is evident from Fig. 6b. But it is not used as structurally; it is difficult to mount and rotate on a motor.

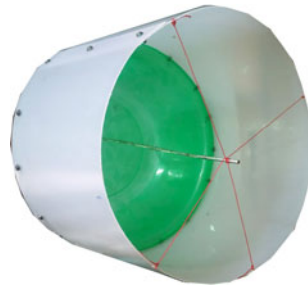


Fig. 5 Directivity comparison between unidirectional microphone without the reflector; parabolic reflector only and cylindrical mouth parabolic reflector

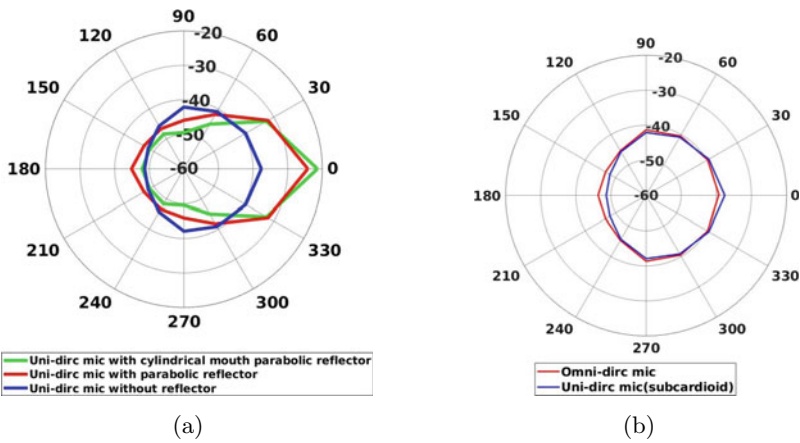


Fig. 6 **a** Directivity comparison between unidirectional microphone without the reflector; parabolic reflector only and cylindrical mouth parabolic reflector **b** Directivity comparison between omni and unidirectional microphone

3.3 Directionality and Source Frequency

By the previous experiment, we observed that the mosque prayer sound did not show any directivity. The reason for this could be that the source is far and the feeble scattered sound reaches the reflector from all directions. Also, the sound becomes low pass through the long-distance propagation through the medium. To understand the possible frequency selectivity, we set up an experiment in which we played different frequencies at different distances and recorded the same through the rotating parabolic reflector. We have calibrated the mic speaker setup. The calibration is required since both the mic and speaker themselves have certain frequency selective responses and efficiency. Calibration was done by keeping the mic and speaker at a 0.5 m distance, and the signal level is adjusted to 98 dB received level.

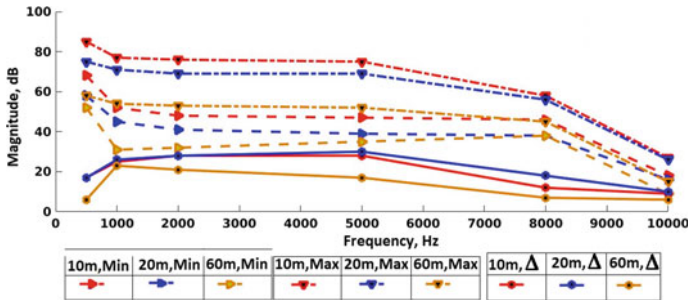
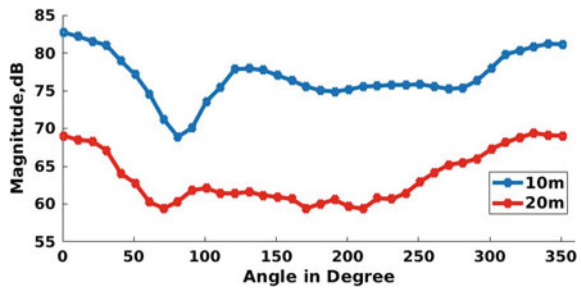


Fig. 7 Change in signal energy with distance and frequency (outdoor recording)

Fig. 8 Energy of signal 500 Hz at different distance



The recording was done for 10, 20 and 60 m source distances and 0.5 kHz, 1 kHz, 2 kHz, 8 kHz and 10 kHz frequency signals.

We calculate the STFT so that we can calculate the energy of a given frequency. For the calculation, we have used the following parameter. Hamming window was used as a window function. $wlen = \text{Window Length} = 1024$; Hop size = $\frac{1024}{4} = 256$ The FFT is of 2048. Consider S_f as the FFT of the frame. The energy was given by S.

$$S = 20\log_{10} \frac{|S_f|}{wlen}$$

The output of the directionality experiment 500 Hz at 10 and 20 m distance is shown in Fig. 8. The inference of which is drawn in Fig. 7 for all the frequencies and distances which shows that lower frequency sound has lesser directionality(Δ) and so also higher frequencies. But the magnitude of the lower frequency sound is predominately high for all distances of the source. This is because low-frequency waves travel further than high-frequency waves since there is less energy absorption by the medium. The reason for less directivity is because the gain of the parabolic reflector is given by

$$G = n \frac{4\pi A f^2}{c^2}$$

. Thus, the relative gain provided by the reflector for lower frequency is much less. For farther sources, the signals which can reach the mic are predominantly low pass in nature, and the parabolic reflector provides minimal gain or directional advantage.

Thus, as the distance increases, the energy of the lower frequency signal is more, but due to the parabolic reflector property, the gain by the reflector is less. Which can be prominently seen by Δ (Difference in max and min signal energy) in Fig. 7.

3.4 Audio Data Recording

Common events of interest in near outdoor scenario (of our building) are prayer, car sounds and bird sounds. We have collected different bird sounds and considered them all as one class and also car horn, and car driving sounds are taken as one class (our goal is not to expand on the variety of audio events but rather to see the intelligence that can be gathered based on the time and directionality of the audio event). The sample events chosen are classified hierarchically using different feature measurements . Each event has its characteristic time–frequency pattern. The prayer event has dominant frequency 1500Hz and above which it is mostly environmental background. A spectrogram of different events is shown in Fig. 9. Table 3 describes the total data that is used to train models at each level of the classification. Each audio training data clip size is of 1 s. We have annotated the data manually using “praat” and “audacity.” For testing purpose, we have used a continuous 4 h of annotated audio data, where each event has over 10 min of data.

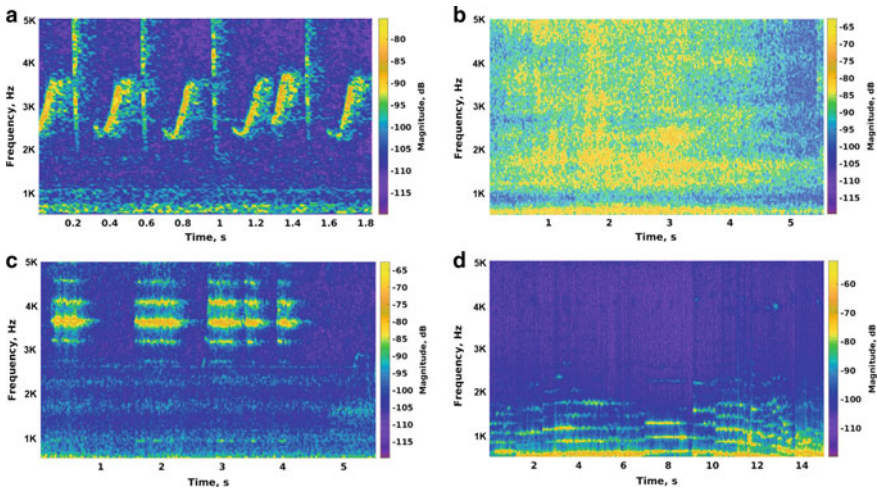


Fig. 9 Spectrogram of different events a bird sound b car sound c prayer d car horn

Table 3 Training data distribution for different models and model parameters

Training data			GMM mix	Frequency band (Hz)
Models	Total files	Total duration		
Prayer and non-prayer	1200 and 1200 (600 car and 600 birds)	20 min and 20 min	4	500–1500
Car and bird	1200 & 1200	20 min and 20 min	8	1500–5000

Table 4 Event classification accuracy

S. No.	Events	Accuracy (%)
1.	Bird	94.88
2.	Car	95.6
3.	Prayer	91.56

4 Audio Event Detection Scheme

The scheme comprises of two parts: offline training and testing of estimated event models and their classification accuracy; online testing for event detection and classification in continuous audio streams.

4.1 Pre-processing and Feature Extraction

We have recorded audio data at different times of the day in the near-outdoor environment and collected a large number of audio clips for each of the events. From the spectrograms of the various events, we observed that the required events are in the frequency range 500–5000 Hz, and the lower frequency band has the unavoidable motor noise and other electrical noise. Hence, we have bandpass filtered the region of interest and then performed energy-based thresholding to detect an event from the background. The energy is calculated over 100 ms frames with 50% overlap. When a frame crosses the threshold, the next 20 frames are retained as likely event frames, since the minimum time for an event is considered to be 1 s. If the event duration is < 1 s, then some of the background noise gets included in the event (may cause false positive). If event duration > 1 s, the whole event will be considered since we are evaluating all frames sequentially. 39 dimensional MFCC features (12 cepstral + 0^{th} coefficient + $\Delta + \Delta\Delta$) are used as parameters of the feature vectors for event modeling and detection.

4.2 Event Modeling

The audio events considered here are as follows: (i) distant mosque prayer, (ii) street car sound and (iii) bird sound. Hierarchical (top-down) approach is used to detect the events. Hierarchical decision is preferred because certain events are more difficult to classify, and they also occur for different durations. We can use different features suitable for different classes. Hence, we expect better results compared to single-level multi-class classification. The events considered here have a common frequency band, and the hierarchy is formed by considering dominant frequency bands for the particular event. Since events properties do vary from instance to instance, we consider stochastic modeling of the event using Gaussian mixture models (GMMs).

GM models are trained at each level of the hierarchy. In the second level, a four-mixture GMM is used to build the universal background model (UBM) using prayer and non-prayer data. For the third level, a eight-mixture GMM is used to build car sound and bird sound models, respectively. The number of mixtures of the model was determined experimentally for optimum performance (more mixtures for greater variability). In this GMM-UBM system, we derive the event-dependent model by adapting the event-specific data through the maximum a posteriori (MAP) adaptation. Totally four models have been trained, prayer, non-prayer, car and bird, respectively, through MAP adaptation based on different feature vectors.

4.3 Detection

All the trained models are used in the detection process. Given the model, the log-likelihood of the test data vectors is calculated. At level-2, the models for comparison will be between prayer and non-prayer. If it is a non-prayer event, then we will go to level-3 for final classification as bird sound or car sound. Maximum log-likelihood score is used in the decision-making at each level.

5 Experimental Results

We have set up the experimental facility on the terrace of the building and observed several hours of audio data recording, and we examined the hierarchical audio detection/classification scheme, both in offline and online mode of operation. The energy threshold for event/background separation is estimated using 1 h of recorded and annotated environmental background data. We have collected background data at different times of the day and have estimated a conservative global threshold not to miss out an audio event (false positive is better than false negative). The event/background decision efficiency is shown in Table 5. Event-specific TP and FN are shown in Table 6.

Table 5 Training data distribution for different models and model parameters

Total duration (min)	Background duration (min)	Vent duration (min)	TP (Event as event)	FN (Event as back-ground)	FP (Back-ground as event)	TN (Back-ground as back-ground)
61.24	30.88	30.66	84.67%	15.33%	13.76%	86.24%

TP True positive, TN True negative, FP False positive, FN False negative

Table 6 Event-specific accuracy

S. No.	Events	Total duration (min)	TP (%)	FN (%)
1.	Car	15.19	83.63	16.37
2.	Bird	6.96	74.71	25.29
3.	Prayer	8.21	95.05	4.95

At level-2, the events have distinguishable frequency characteristic as shown by the spectrogram of Fig. 9. Using MFCC features for each frame of 30 ms and overlap of 10 ms, the event-specific models are trained as shown Fig. 13.

5.1 Testing

Offline In offline testing, the system accuracy is tested on 4 h of manually annotated data. The test data is first segmented into 1 s duration clips, and the decision is determined on each segment. The 4 h of data contained 40 min of car, 10 min of bird and 15 min of prayer data. The testing procedure is shown in Fig. 10. Results are shown in Table 4.

Online For real-time event detection and classification, continuous data is streamed and processed. Because of the data acquisition and computation on laptop, approx 4 min of delay is there to display the events after its occurrence. We store continuous data stream after 60 s and segment it to 1 s clips of audio data and then perform detection decision. Event visualization shows 60 s long audio data, and then, it updates for the next file. Log-likelihood of two hierarchical levels are plotted in Fig. 11 for 60 s data. For clarity of visualization, we have taken a 5 s group and showed the most occurred event over that duration in Fig. 12.

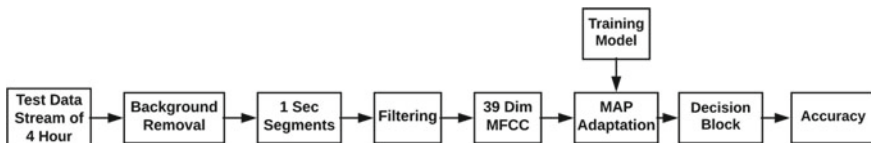


Fig. 10 Offline testing procedure

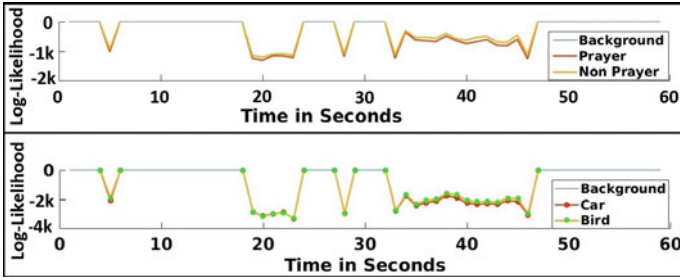


Fig. 11 Log-likelihood is calculated for a 1 min segment of audio which has an overlap of bird and car event, which is accurately predicted by the plot

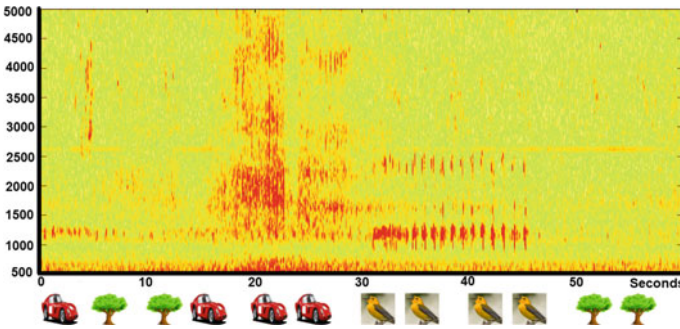
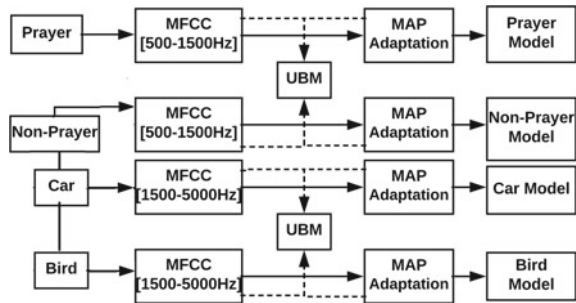


Fig. 12 Spectrogram and event visualization

Fig. 13 Training procedure



5.2 Event Direction

After detecting/classifying the events, we have tried to estimate the direction of the movement of the event by approximating the azimuthal angle of the event. In Fig. 14, we can see a continuous bird event for the full rotation of the reflector. The max. change of sensitivity approximates the direction of the event. From Fig. 14, we can infer that the bird is at 0° to the initial position of the reflector and the difference in

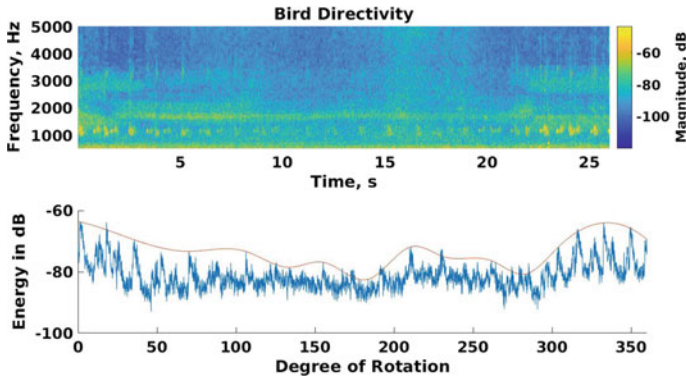


Fig. 14 Event direction

measured sensitivity $\Delta \approx 16$ dB. We can also estimate that the source is about 30–40 meter distance from the microphone, based on the reflector sensitivity that we have earlier calculated.

5.3 Inference

After detecting the occurrence of events, their occurrence rate and their direction, timing, etc. can be collated, which can lead to further higher-level information. Depending on the movement of the cars and the occurrence of bird sounds, we can infer the time of the day (day/night) and which road side is busier. The prayer event usually has a fixed timing depending on the geographical reason, so by detecting prayer, we can comprehend the time of the day. Increasing the variety of events that can be detected, we can gather finer information, which can help us toward finer intelligence around the premises.

6 Conclusion and Future Work

Audio surveillance using rotating reflector mic scheme is promising as a low-cost solution for near-outdoor application. The issues in full-outdoor (large area) or full-indoor (fragmented space) surveillance call for different type of audio sensor design and intelligence gathering solution.

Acknowledgements We thank Prof Thippur V. Sreenivas ,SAG Laboratory , IISc Bangalore for guiding us for the project. We also thank Major Omkar Sawant and Meghna Ranjit for the help in the hardware set up and their guidance. Without them, many of the experiments would not have been possible.

References

1. Y. Nishida, T. Hori, T. Suehiro, S. Hirai, Monitoring of breath sound under daily environment by ceiling dome microphone, in *2000 IEEE International Conference on Systems, Man, and Cybernetics*, vol. 3. (IEEE, 2000), pp. 1822–1829
2. H.M. Aumann, T. Russell, N.W. Emanetoglu, Comparison of a small parabolic reflector for use with an acoustic and a radar microphone, in *IEEE International Symposium on Antennas and Propagation & USNC, URSI National Radio Science Meeting* (IEEE, 2015), pp. 2199–2200
3. A. Temko, C. Nadeu, Acoustic event detection in meeting-room environments. *Pattern Recogn. Lett.* **30**(14), 1281–1288 (2009)
4. J. Schröder, S. Goetze, J. Anemüller, Spectro-temporal gabor filterbank features for acoustic event detection. *IEEE/ACM Trans. Audio, Speech, Language Proc.* **23**(12), 2198–2208 (2015)
5. P.K. Atrey, N.C. Maddage, M.S. Kankanhalli, Audio based event detection for multimedia surveillance, in *2006 IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings*, vol. 5 (IEEE, 2006) , pp. V–V
6. J. Li, W. Dai, F. Metze, S. Qu, S. Das, A comparison of deep learning methods for environmental sound detection, in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2017), pp. 126–130
7. M. Crocco, M. Cristani, A. Trucco, V. Murino, Audio surveillance: a systematic review. *ACM Comput. Surveys (CSUR)* **48**(4), 52 (2016)
8. S. Moncrieff, S. Venkatesh, G. West, Online audio background determination for complex audio environments. *ACM Trans. Multimedia Comput. Commun. Appl. (TOMM)* **3**(2), 8 (2007)

Performance Evaluation and Synthesis of FIR Filters Using Various Multipliers Algorithms



Aniket Kumar  and R. P. Agarwal

Abstract With the advancement of sensors and network-enabled devices, the need for high performance and low power consumption is increasing day by day. The major delaying part of these devices is the multipliers built into the digital filters for performing signal processing applications. This paper proposes a design of various conventional and traditional multiplier algorithms to design finite impulse response (FIR) filter of order four, eight, sixteen, thirty two and sixty four and its performance analysis in terms of delay, memory usage and level of logic used.

Keywords FIR digital filter · Delay · Filter design and analysis (FDA) · Finite state machine (FSM)

1 Introduction

Nature is surrounded by numerous types of signals in the form of human voice, audio signals, seismic waves, etc., that are used in different applications of signal processing and communications. Signal processing has a rich history in the diverse fields as speech communication, biomedical engineering, sonar, radar, seismology, data communication and many more. In the multitude of diverse fields of science and technology, digital signal processing (DSP) has become an important modern tool, converts analog input signal to digital, that is processed digitally, and then converts back to analog signal [1].

With the development of sophisticated signal processing algorithms, recent advancements in integrated circuit technology led to the development of DSP systems. Finite impulse response (FIR) filtering is one of the fundamental steps in many DSP applications, and to optimize it in terms of delay, area, LUTs and level of logic, conventional as well as traditional multiplier algorithms that are used to design FIR filters are implemented, and their findings are used to implement FIR filter of order 4, 8, 16, 32 and 64. Complexity in multiplier design controls response

A. Kumar (✉) · R. P. Agarwal

Department of Electronics and Communication, Shobhit Institute of Engineering and Technology (Deemed-to-be-University), Meerut, Uttar Pradesh 250110, India

time of filter, therefore to optimize its design, is an important task that has been focused in many previous works [2–9] in which digital multiplier has been designed for different bit lengths. Sections 2 and 3 give the overview of selected multiplier algorithms’ FIR filters. Section 4 focuses on simulation and implementation results of multiplier algorithms as well as filter design. Finally, the proposed work is concluded in Sect. 5.

2 Multiplier Algorithms

As multiplier plays an import role in designing a filter, many algorithms have been used such as Vedic, Array, Wallace, Dadda, Booth, sequential algorithms and many more having different approaches to find product of two binary number [10, 11].

2.1 Vedic Algorithm

This algorithm uses concepts of ancient Vedic sutra “Urdhva–Tiryabhyam,” a technique that uses vertical and crosswise technique step-wise to find product of numbers as discussed in Fig. 1 [7, 12].

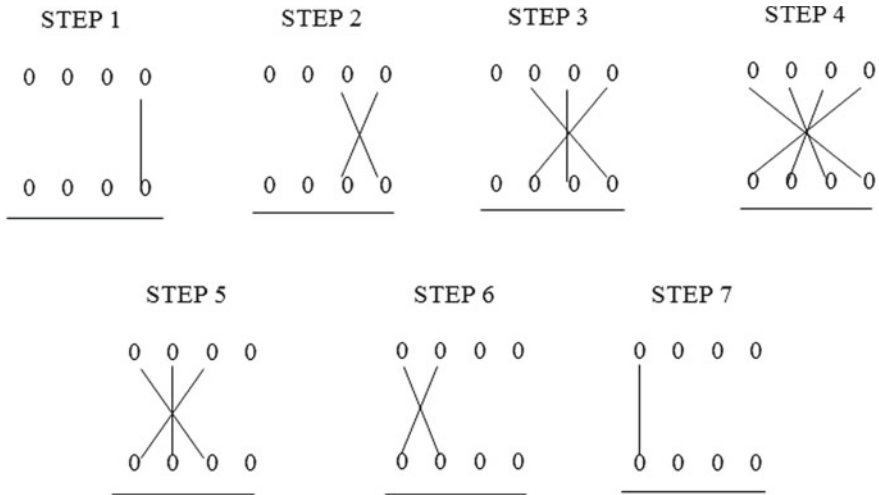


Fig. 1 Line diagram showing multiplication of two binary numbers [2]

2.2 *Array Algorithm*

This algorithm is based on the sequence of addition and shifting procedure to obtain partial products generated step-wise that are arranged in different rows which are added column-wise to obtain final product terms [3].

2.3 *Wallace Algorithm*

To evaluate the product of large number, this algorithm is used where partial product terms are arranged in a tree-like format and reduces critical path delay by using full adder as 3:2 compressor and half adder as 2:2 compressor [4].

2.4 *Dadda Algorithm*

Also, in this algorithm, partial products are arranged in a tree-like format, and the reduction technique used is governed by the formula $d_{j+1} = \text{floor}(1.5d_j)$ where d_j is a maximum height of the column in a tree-like format [5].

2.5 *Booth Algorithm*

In Booth algorithm, partial product and their addition are done by encoding technique [13, 14]. Multiplicand bits are grouped from MSB to LSB, and as per encoding technique, operations such as addition, subtraction and shifting are done to obtain the final result [5, 6, 15].

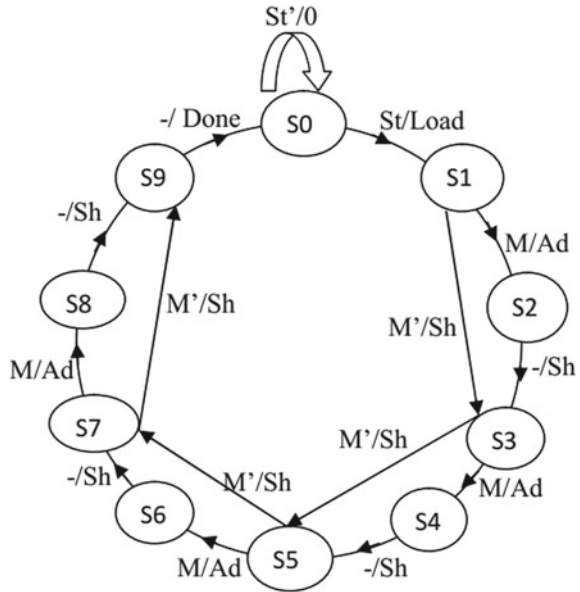
2.6 *Sequential Algorithm*

Sequential algorithm uses shift and add method to obtain the final result as explained by Fig. 2 showing finite state machine (FSM).

3 **FIR Digital Filter**

To meet high speed requirement and stability in various applications of digital filters used, finite impulse response (FIR) filter is preferred over infinite impulse

Fig. 2 Sequential algorithm (FSM) [7, 8]



response (IIR) filter. Due to linear phase, stability and low complexity, FIR filter is used in multimedia communication and mobile communication systems. Many previous works [9, 16, 17] have focused on its designed; however, in this manuscript, Equiripple design method has been used for low-pass filter for order four, eight, sixteen, thirty two and sixty four.

4 Simulation and Implementation Results

VHDL codes of all the mentioned multiplier algorithm are simulated on MODELSIM and implemented on FPGA using Xilinx software for bit length two, four, eight, sixteen, thirty two and sixty four to obtain analysis report in terms delay, hardware complexity, LUTs and IO parameters [7–9, 11].

Table 1 shows an analysis report that is obtained after simulated of VHDL code of an algorithm for a particular bit length.

Comparative graphs are plotted using MATLAB 2019b tool for delay and memory usage parameters. Figure 3 shows delay in response concerned to the above said algorithms, and Fig. 4 shows memory consumed in the said algorithms to produce final product terms.

As sequential algorithm has higher mean deviations of delay and due to that for comparative delay graph, sequential plot has been ignored as shown in Fig. 3, and graph in black indicates average delay. Figure 4 shows a comparative analysis of memory (KB) used by used multiplier algorithms on Spartan 6.

Table 1 Compression of multiplier algorithms for different bit lengths

Bit length	Type	Delay in ns	Levels of logic	Number of slice LUTs	Memory in KB
2	Vedic	5.505	3	4	259,592
	Array	5.505	3	4	259,592
	Wallace	6.494	4	6	258,052
	Dadda	6.494	4	6	258,052
	Booth	2.606	2	5	259,080
	Sequential	24.68	2	15	258,312
4	Vedic	9.557	7	23	246,848
	Array	10.697	8	23	246,848
	Wallace	17.00	7	23	322,176
	Dadda	14.02	7	19	246,144
	Booth	13.548	12	57	247,872
	Sequential	46.26	2	26	320,832
8	Vedic	18.270	15	121	253,312
	Array	22.055	19	99	354,816
	Wallace	19.829	15	128	283,648
	Dadda	13.907	11	114	282,176
	Booth	19.081	28	235	249,216
	Sequential	83.104	2	37	243,584
16	Vedic	27.278	23	656	284,608
	Array	40.341	37	387	359,296
	Wallace	36.731	29	499	361,600
	Dadda	22.124	19	453	359,104
	Booth	27.742	67	834	252,288
	Sequential	189.585	2	90	252,288
32	Vedic	47.146	41	2120	339,648
	Array	78.313	74	1539	414,336
	Wallace	68.383	57	2147	342,272
	Dadda	46.720	43	1716	313,984
	Booth	51.826	131	3202	299,840
	Sequential	375.83	3	150	183,564
64	Vedic	80.146	73	8793	378,560
	Array	174.583	71	6744	548,480
	Wallace	143.10	103	8531	393,664
	Dadda	149.148	80	7023	360,064
	Booth	194.653	454	25,221	360,960
	Sequential	887.391	2	233	183,579

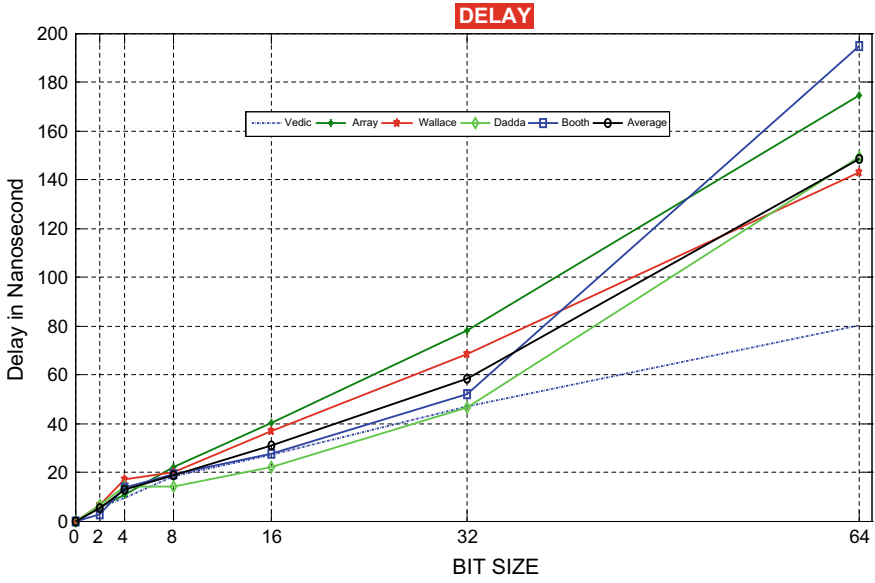


Fig. 3 Comparative delay graphs for selected multiplier algorithms

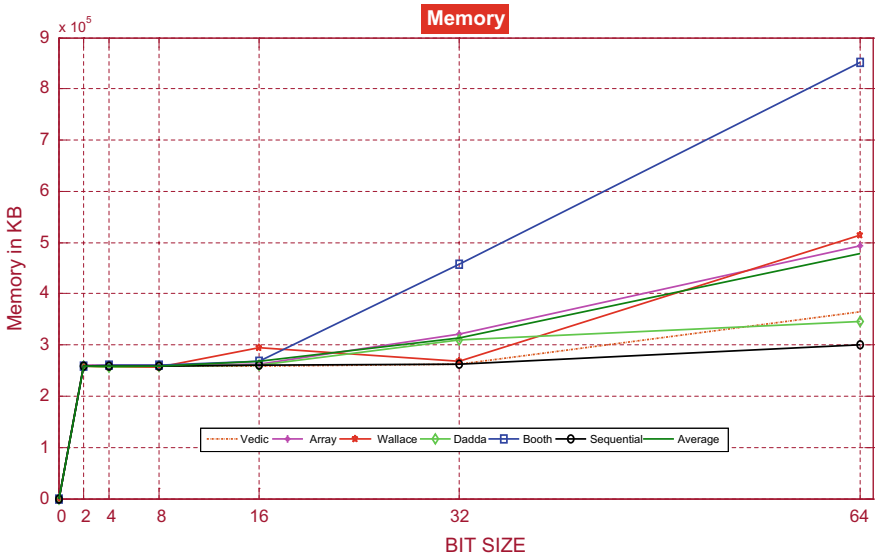


Fig. 4 Comparative memory usage graphs for selected multiplier algorithms

Using the above said algorithms as a component, low-pass FIR filter is coded in VHDL, simulated and implemented on FPGA using Xilinx software for filter order four, eight, sixteen, thirty two and sixty four. Table 2 shows an analysis report of FIR filter obtained from analysis report in terms delay, hardware complexity, LUTs and IO parameters.

Table 2 Comparison of FIR filter of order 4,8,16, 32 and 64 using Wallace, Dadda, Booth, sequential and array algorithm on Device Spartan6, XC6SLX45-CSG324

Filter order	Variant	Delay in ns	Levels of logic	Number of slice LUTs	Memory in KB
4	Vedic	9.236	1	921	246,848
	Array	9.997	1	1004	246,848
	Wallace	15.255	19	849	322,176
	Dadda	7.363	12	773	246,144
	Booth	8.328	1	863	247,872
	Sequential	20.328	8	789	320,832
8	Vedic	11.267	20	1451	253,312
	Array	11.276	21	1273	354,816
	Wallace	12.003	24	1826	283,648
	Dadda	11.102	20	1575	282,176
	Booth	12.036	12	115	249,216
	Sequential	90.136	27	236	243,584
16	Vedic	15.626	29	3199	284,608
	Array	16.037	31	2716	359,296
	Wallace	15.926	28	3698	361,600
	Dadda	15.843	30	3378	359,104
	Booth	16.856	20	2834	252,288
	Sequential	250.016	68	2290	252,288
32	Vedic	21.255	46	5970	339,648
	Array	22.482	40	5882	414,336
	Wallace	20.802	39	7657	342,272
	Dadda	22.494	43	6342	313,984
	Booth	23.405	45	3205	299,840
	Sequential	665.664	70	4150	183,564
64	Vedic	29.117	78	11,558	378,560
	Array	34.948	56	12,187	548,480
	Wallace	29.096	58	15,981	393,664
	Dadda	34.848	80	13,644	360,064
	Booth	35.088	71	17,023	360,960
	Sequential	1862.144	73	16,744	183,579

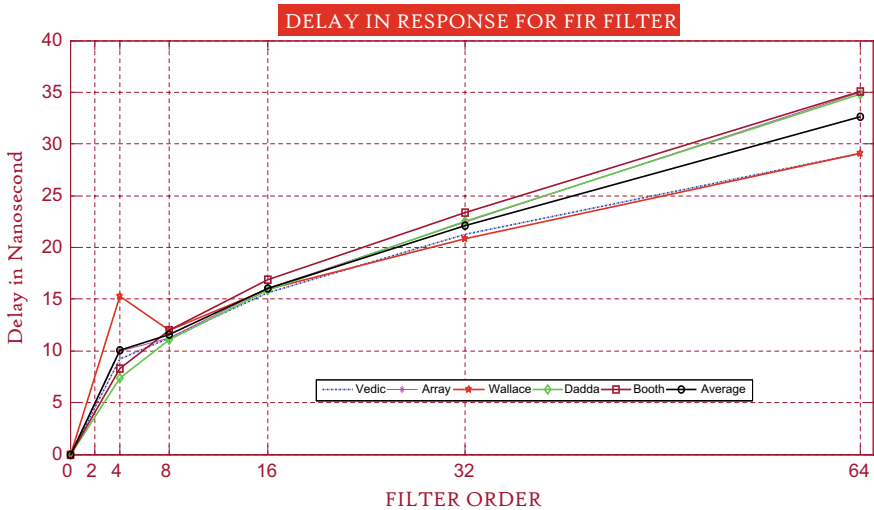


Fig. 5 Comparative delay graphs for selected multiplier algorithms

Coefficients of direct form low-pass FIR filter are obtained from filter design analysis (FDA) tools using Equiripple method of design [18–20]. Comparative graphs are plotted using MATLAB 2019b tool for delay and memory usage parameters. Figure 5 shows delay in response FIR filter for different orders and Fig. 4 shows memory consumed in the said filter order to produce the desired output.

From analysis report shown in Table 2, delay for sequential block when it is used for filter design has maximum mean deviation; therefore, while plotting comparative graph, sequential block technique has been ignored, and in the same graph, average delay has been calculated and plotted for selected filter order.

A Comparative analysis of memory (KB) consumed by the particular algorithm block to design low-pass FIR filter for filter order of four, eight, sixteen, thirty two and sixty four on Spartan 6 as shown in Fig. 6.

5 Conclusion

From simulation, implementation, analysis report (Table 1) and comparative graphs of multiplier algorithms (Figs. 3 and 4), it is concluded that Vedic algorithm is the best choice in terms of speed for high-bit-sized, as its delay is 80.146 ns followed by Wallace with approximately twice delay of 143.10 ns followed by Dadda algorithm with approximately ten times delay in sequential algorithm.

Analysis of multiplier algorithm alone and then using them in an application like FIR filter design we concluded (from Tables 1 and 2, Figs. 5 and 6) that, in terms of delay, Vedic algorithm is the best when it is used alone as multiplier, however

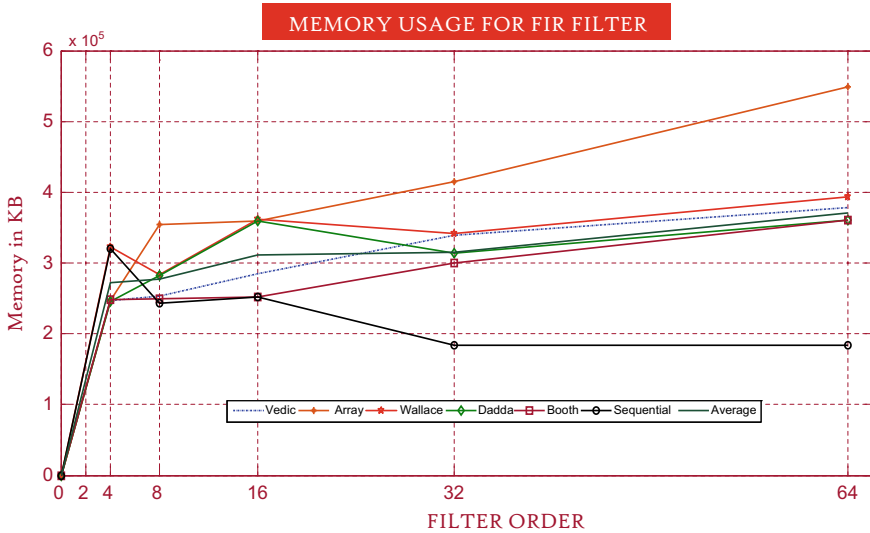


Fig. 6 Comparative delay graphs for selected multiplier algorithms

when selected proven multiplier techniques are used for FIR digital filter design, it is deduced that Wallace algorithm for high order low-pass filter is best in terms of delay, followed by the low-pass FIR filter based on Vedic algorithm with minor difference. However, with mean deviation of delay and memory usage, the Vedic Logic leads over all the proved multiplier algorithms.

References

1. G.G. Kumar, S.K. Sahoo, Implementation of a high speed multiplier for high performance and low power applications, in *Proceedings of the IEEE 19th International Symposium on VLSI Design and Test (VDAT) (2015)*, pp. 1–4
2. V. Sharma, A. Kumar, Design, implementation & performance of vedic multiplier for different bit lengths. *Int. J. Innovative Res. Comput. Commun. Eng.* **5(4)**, 7912–7919 (2017). ISSN: 2320-9801(O), 2320-9798(P)
3. A. Kumar, V. Sharma, Comparative analysis of vedic & array multiplier. *Int. J. Electron. Commun. Eng. Technol.* **8(3)**, 17–27 (2017). ISSN: 0976-6464 (P), 0976-6472
4. A. Kumar, E. Gupta, R.A. Shobhit, R.K. Jain, Comparative research for managing delay in signal processing via multipliers, in *Proceedings of 2nd IEEE International Conference on Power Electronics, Intelligent control and energy Systems (ICPIECES-2018)*, DTU, India (2018), pp. 1549–1554. ISBN: 978-1-5386-6626-5
5. A. Kumar, R.P. Agarwal, An approach to manage delay in signal processing via selected multiplier algorithms, in *Proceedings of 3rd IEEE International Conference on Inventive Computation Technologies (3rd ICICT 2018) Conference Proceedings*, RVS Technical Campus, Coimbatore, Tamil Nadu, India (2018), pp. 401–405. ISBN: 978-1-5386-4985-5

6. A. Kumar, R.P. Agarwal, Simulation & implementation of efficient multiplier circuits, in *Proceedings of 2nd International Conference on Recent Multidisciplinary Research (ICRMR-2018)*, Asian Institute of Technology Conference Centre, Thailand, 23–24 November 2018, pp. 1–6. Published in Online Int. Interdisc. Res. J. (OIIRJ) **8**(2), 104–111 (2018). ISSN: 2249-9598. UGC J. No. 46964
7. A. Kumar, R.P. Agarwal, Complex multiplier: implementation using efficient algorithms for signal processing application. *Int. J. Recent Technol. Eng.* **8**, 1235–1239 (2019). ISSN: 2277-3878, UGC J. No. 49239
8. A. Kumar, R.P. Agarwal, Delay estimation, area analyses and comparison of digital multiplier techniques for signal processing applications, in *Proceedings of International Webinar (e-conference) on “Prospective of Interdisciplinary Research in Science and Technology in Present Scenario”*, Department of Physics, CCSU, Meerut (2020)
9. D. Behera, A. Patnaik, P.K. Barik, G.S. Rath, VLSI implementation of digital filter using novel RTSD adder and booth multiplier. *Int. J. Eng. Adv. Technol.* **9**, 4131–4139 (2020)
10. R. Shetty, M.B. Neelagar, Design and implementation of high performance 4-bit dadda multiplier using compressor. *Int. J. Comput. Sci. Mobile Comput.* **6**(7), 249–254 (2017)
11. P.K. Sinha, P. Sharan, multiplexer based multiplications for signal processing applications. *Indonesian J. Electr. Eng. Comput. Sci.* **9**(3), 583–586 (2018)
12. A.S. Thakur, V. Tiwari, Design high speed FIR filter based on complex vedic multiplier using CBL adder, in *2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE)*, Bhubaneswar, India (2018), pp. 559–563
13. M. Tiwari, A. Kumar, Implementation of high speed and low power novel radix 2 booth multiplier using 2248 BEC converter. *Int. J. Eng. Sci. Comput.* **7**, 4861–4863 (2017). ISSN 2321 3361
14. A. Kumar, M. Tiwari, Simulation, implementation & comparative analysis of multiplier, in *Proceedings of International Conference on Recent Trends in Operations Research & Statistics (RTORS-2017)*, IIT Roorkee, India (2017), pp. 29, 28–30
15. E. Gupta, A. Kumar, R.K. Jain, Simulation & comparative analysis of booth multiplier. *Int. J. Electron. Eng.* **10**(2), 404–412 (2018). ISSN: 0973-7383, UGC J. No. 2946
16. A. Mittal, A. Nandi, D. Yadav, Comparative study of 16-order FIR filter design using different multiplication techniques. *IET Circ. Devices Syst.* **11**, 196–200 (2017)
17. Mamta, A. Kumar, Comparative analysis of FIR filter. *Int. J. Adv. Eng. Res. Dev.* **3**, 88–95 (2016). ISSN 2348-4470(O), 2348-6406(P)
18. S. Soundarya, N. Arumugam, R. Rameshkumar, Design of FIR filter using efficient adder and multiplier for ECG signal processing applications. *Int. J. Eng. Appl. Sci. Technol.* **2020**(4), 276–282 (2020)
19. R.A. Sekar, S. Sasipriya, Implementation of FIR filter using reversible modified carry select adder, in *Special Issue: Recent Research Challenges in Intelligent Techniques, Big Data and Internet of Things*, vol. 31 (2019), pp. 1–9
20. A. Kumar, Mamta, Comparison of different types of IIR filters. *Int. J. Adv. Res. Electron. Commun. Eng.* **5**(2), 393–402 (2016). ISSN: 2278-909X

Design and Implementation of Bluetooth Low Energy Link Layer Controller Using Dataflow Programming



Nur Atiqah Abdul Latib, Ab Al-Hadi Ab Rahman,
and Mohd Shahrizal Rusli

Abstract This paper presents the design of Bluetooth low energy (BLE) link layer controller implemented on application-specific integrated circuit (ASIC), field-programmable gate array (FPGA) and general purpose processor (GPP). CAL dataflow programming has been used that allows automatic RTL and C code generation from high-level specification. The generated RTL code is synthesized and implemented on ASIC Silterra 180 nm process technology and Xilinx Artix FPGA, while the generated C code is implemented on Intel i7 GPP. High-level design space exploration has been made by performing actor merging techniques to obtain seven different architectures, which are compared in terms of performance, area, power and energy. The results show that energy improvements for ASIC, FPGA and GPP are 44.71%, 10.22% and 28.31%, respectively, through the design space exploration methodology.

Keywords Bluetooth low energy · Dataflow programming · ASIC · FPGA · Design exploration

1 Introduction

Bluetooth was introduced to replace serial communication interfaces that enable devices to communicate wirelessly. It emphasized on low-cost, low-power and short-range transmission between devices [1]. Nowadays, Bluetooth is widely used in most wireless devices such as mouse, keyboards, earphone, printers and others.

N. A. A. Latib (✉) · A. A.-H. A. Rahman · M. S. Rusli
School of Electrical Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Malaysia

A. A.-H. A. Rahman
e-mail: hadi@utm.my

M. S. Rusli
e-mail: shahrizal@utm.my

To be precise, these devices are designed with Bluetooth low energy (BLE) [2] features to be used with longer time period and low power consumption to conserve the battery life.

The architecture of BLE consists of three main blocks which are application, host and controller. Controller block consists of three sub-blocks which are link layer, physical layer and direct test mode. The controller and the host communicate through the host controller interface (HCI). This paper focuses on the link layer controller that controls the process of transmitting and receiving data between devices.

However, the challenge in designing the controller is the complexity [1]. It is responsible for many tasks that are required by the Bluetooth application such as advertising, scanning, creating and maintaining connections and processing the packet contents.

To overcome these problems, this paper proposes to use the CAL actor language to model the BLE controller. CAL actor language is a high-level dataflow programming that is suitable for designing a complex system [3], with the ability to automatically synthesize its code to RTL for hardware implementation [4], and C for software implementation [5]. To ensure the design meets the Bluetooth specification, high-level actor merging optimization techniques [3] are also applied.

Some of the related works in the design and implementation of the BLE are reviewed as follows. The work in [6] was based on Bluetooth version 1.1 which is the older version of Bluetooth. Thus, the controller is not actually implemented. But some of the functions are slightly similar with BLE functions. In [7], a controller for a specific serial bus which supports high speed, hot plug play and auto-reconfiguration is called IEEE 1394. This literature referred to the Bluetooth version 4.0. This work however does not have a numerical result to be measured and compared.

The recent research that is done in [8] is based on the BLE version 4.1. The research focuses on the whole controller of the BLE. In their research, an intermediate layer is introduced to implement some controller function using the software stack. However, only top-level architecture of the link layer was proposed without any actual implementations.

The work in [9] implemented the BLE for building Internet of things (IoT). The research focused on the host side to foster the IoT innovation by presented open hardware for BLE architecture. They also provide an open-source Contiki operating system (OS) port for new hardware. However, the approach implemented the controller into a microcontroller instead of a custom hardware.

BLE also has been widely implemented on different groups of application such as home automation [10], health care [11] and smart cities [12]. Other than that some of the works in hardware implementation of communication systems are given in [13] and [14]. Some of the works that use the CAL actor language in design and implementation include [15] and [16]. There is also a work that uses CAL actor language to model the ZigBee baseband transmitter [4]. CAL implementation of the BLE controller has not been reported in the literature.

2 Design Methodology

2.1 State Machine from Message Sequence Chart

Based on the message sequence charts (MSC) in the BLE specification, internal state machine for each state can be extracted. Figure 1 shows one example of MSCs that illustrate the handshaking process for initiating connection between two BLE devices.

In Fig. 1, Host A represents the host for Bluetooth Device A, LL A represents the link layer for Bluetooth Device A, LL B represents the link layer for Bluetooth Device B and Host B represents the host for Bluetooth Device B. To initiate a connection with other devices, one device should enter the initiating state (LLA) and one device must enter the advertising state (LLB).

To determine the internal state machine, these steps are followed:

1. Identify the point of receiving and sending data.
2. Determine the type of receiving and sending data.
3. Study the data processing and the state involves.
4. Determine the number of internal states and actions required.
5. Model the link layer using CAL dataflow program.

Based on Fig. 1, five points of receiving and sending data are determined on LL A. As an example, by following the five steps, the information for point one is discovered:

1. Type of data packet: LE create connection (HCI command).
2. Process involve: receiving HCI command from Host A.
3. State involve: initiating state.
4. Number of internal state(s): one.
5. Number of action(s): four.

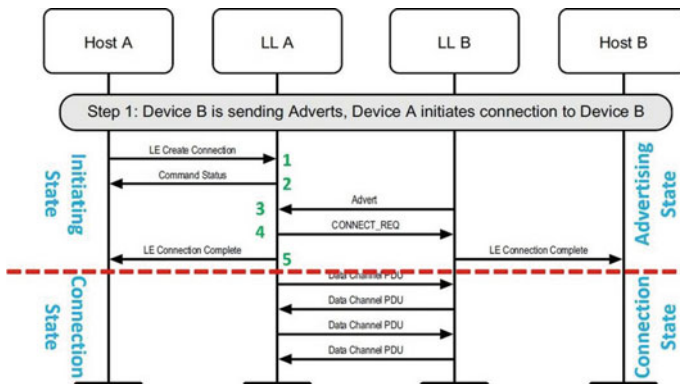


Fig. 1 Example of BLE message sequence chart for states derivation

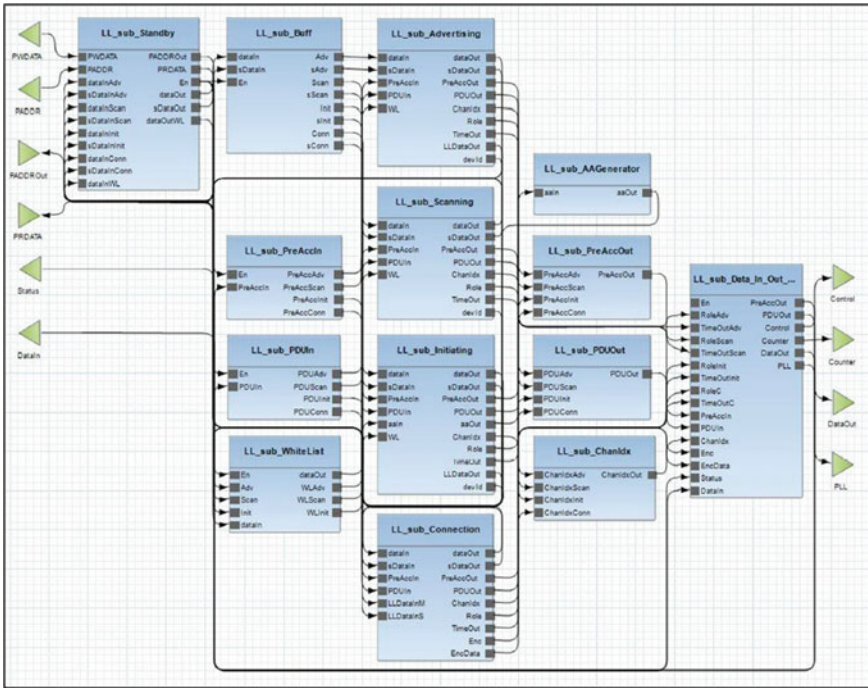


Fig. 2 High-level block diagram of the BLE controller in CAL design environment

2.2 Proposed High-Level Design of BLE Controller

Based on the extracted internal states machine, BLE controller is designed at high level as in Fig. 2. Other than that, the link layer controller is also designed based on the type of packets used, data process mechanism involved, timing requirements, etc. The controller consists of 14 actors which are standby, buff, out. The standby, advertising, scanning, initiating and connection each represent the link layer controller state machine.

2.3 Functional Validation

Before optimization techniques are applied, the functionality of the BLE link layer controller is validated first. In addition, the functional validation is performed after each optimization stage has been done. As mentioned before, the BLE controller is designed at high level. Thus, there are extra steps to validate the designed BLE link layer controller functionality. The validation processes involve validation at high level and simulation waveform. For this work, there are three stages of validation:

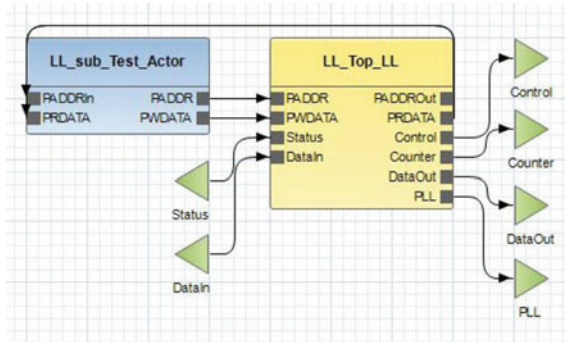


Fig. 3 High-level block diagram to test the BLE controller in CAL design environment

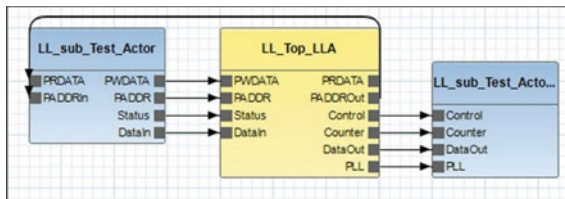


Fig. 4 High-level block diagram to test the BLE controller for the generated C code in CAL

1. High-level validation
2. C code validation
3. RTL testbench validation.

For high-level validation, a test actor is designed to inject input to the link layer controller. The function tested for each validation is different for each test actor. As an example, the connection between test actor and the link layer controller to validate the standby functionality is shown in Fig. 3.

As for C code validation, the method is quite similar with the high-level validation. The only difference is a print actor to print out all the output from the link layer controller. Later, a C code is generated automatically and validated using MinGW tools. The connection between test, link layer controller and print actors is shown in Fig. 4.

On the other hand, for RTL testbench validation, a Verilog code of testbench is designed to validate the generated RTL code of the link layer controller.

2.4 Architectural Optimization

Architectural optimizations are performed using the high-level actor merging techniques as well as the state reduction techniques. Actor merging method is a technique

where two or more actors are combined into one new named actor. To apply this technique, the original actors and a leader actor shall be determined first. Original actors are a set of actors that merged into one new actor while Leader actor is an actor in which the structure of its finite state machine (FSM) is chosen, followed and merged with other actor FSM. Later, state reduction techniques are applied to optimize the design FSM. Figure 5 shows the process of merging actors with comparisons of the number of pins between actors and the number of actors. As a result, the optimized design is obtained, as shown in the simplified block diagram in Fig. 6.

3 Experimental Results

3.1 Performance

The BLE link layer controller is designed to operate at 16 MHz for FPGA and ASIC. This frequency is chosen because it is enough to achieve the maximum data throughput. On the other hand, the frequency is fixed to 2.4 GHz for GPP. With these requirements, the LLG design is used to validate on three different platforms by using the existing five testbenches. Table 1 shows the total time taken for each complete testbench from three different platforms.

Based on Table 1, the FPGA and ASIC have the same total time taken for all testbenches. This is because the design was synthesized on both platforms which were operated at the same frequency of 16 MHz. On the other hand, the GPP operated at a different frequency which is 2.4 GHz. Based on the tabulated data, it is found that the FPGA and ASIC could perform faster than GPP. This can clearly show the differences between the total times taken for each testbench. The standby test bench has the smallest difference, while the advertising testbench has the biggest difference.

3.2 Power Consumption

Other than performance, power consumption is also equally important for hardware design. This is because the power usage will affect the energy used for a device. The power consumption for all seven architectures from all three different platforms is compared. The power consumption recorded in Table 2 is the power obtained at 16 MHz for FPGA and ASIC, while according to specification [16], the power consumption for Intel i7 is fixed to 15 W.

Based on Table 2, it is found that the power consumption on ASIC is higher than FPGA. This is because different technologies used the FPGA used Artix with 14 nm technology, while the ASIC used 180 nm Silterra Process technology. The 14 nm technology for FPGA thus provides less power consumption, and this significant difference is the reason for the ASIC to have higher power consumption compared

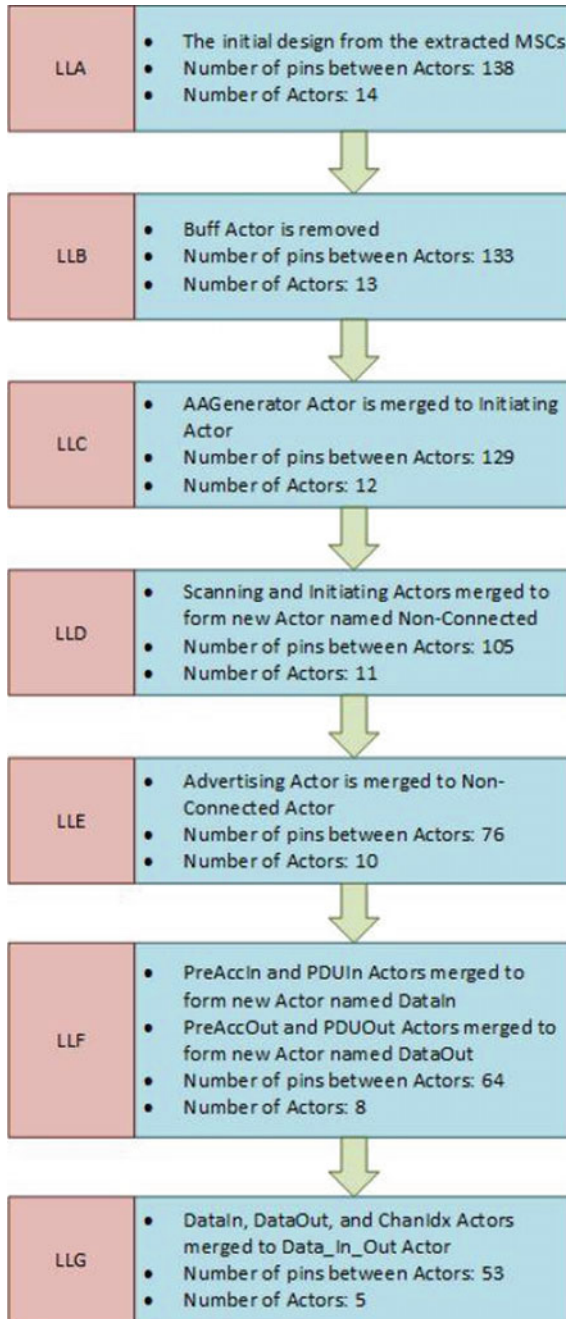


Fig. 5 Applying the actor merging techniques and optimization from initial design (LLA) to final design (LLG)

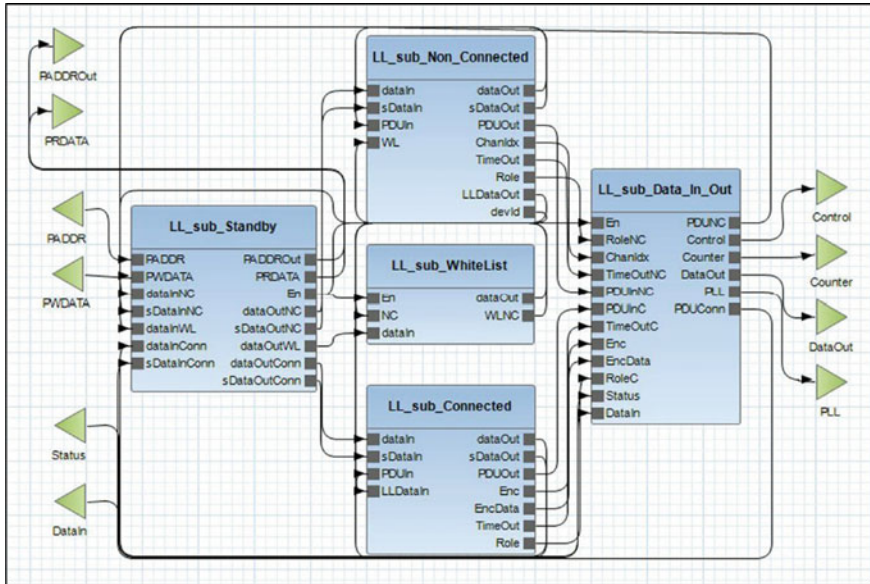


Fig. 6 High-level block diagram for the final LLG design

Table 1 Time taken for each testbench sequence for the three different platforms

Test sequence	FPGA at 16 MHz (ms)	ASIC at 16 MHz (ms)	GPP at 2.4 GHz (ms)
Standby	2.6	2.6	3.0
Advertising	9.3	9.3	14.0
Scanning	6.8	6.8	9.0
Initiating	6.0	6.0	8.0
Initiating cancel	3.4	3.4	4.0

Table 2 Power consumption for all seven architectures for the three different platforms

Design	FPGA (mW)	ASIC (mW)	GPP (mW)
LLA	175	401	15,000
LLB	172	374	15,000
LLC	176	365	15,000
LLD	158	322	15,000
LLE	160	272	15,000
LLF	160	242	15,000
LLG	163	222	15,000

Table 3 Average energy consumption for all seven architectures for the three different platforms

Design	FPGA (mJ)	ASIC (mJ)	GPP (mJ)
LLA	0.9824	2.2526	159
LLB	0.9656	2.1060	147
LLC	0.9880	2.0550	132
LLD	0.8870	1.8128	132
LLE	0.8982	1.5335	114
LLF	0.8982	1.3639	114
LLG	0.9151	1.2468	114

to FPGA. On the contrary, the GPP got highest power consumption, and it was the same for all design architectures.

3.3 Energy Consumption

With acquired power consumption from all architectures that were implemented on three different platforms, the average energy consumption for all design architecture can be calculated. The following equations were used to calculate the average energy consumption: To calculate the energy:

$$E = P \times t \quad (1)$$

where E is energy, P is power and t is the time for a complete testbench. While to calculate the average energy for N test sequences:

$$E_{\text{avg}} = \frac{\sum_{i=0}^N E_i}{N} \quad (2)$$

Using these equations, the average energy consumption for all architectures on three different platforms is given in Table 3.

3.4 Resource Utilization

For this aspect, only FPGA and ASIC implementations were compared. This is because there is incomparable information for resources used for GPP implementation. Each report for resource utilization obtained from Vivado tools is visualized in a graph as shown in Fig. 7.

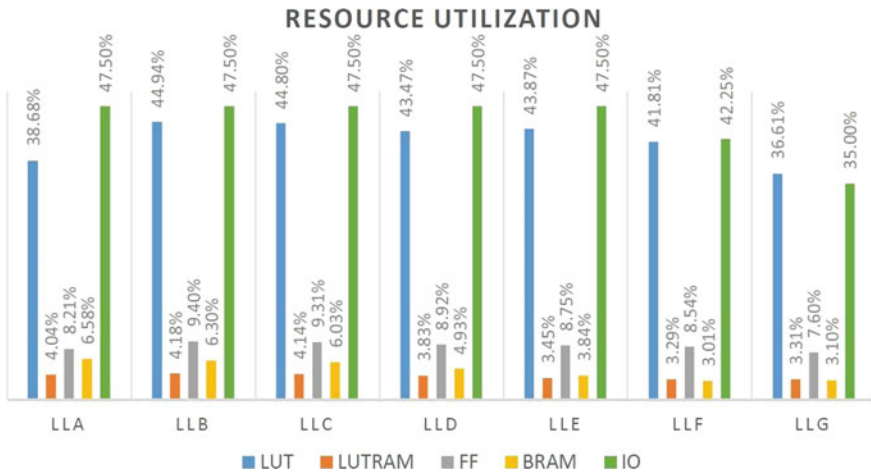


Fig. 7 Resource utilization in FPGA for all seven design architectures

From the graph in Fig. 7, there are four important components to be observed which are LUT, LUTRAM, FF and BRAM because these four are the major components on FPGA. Based on the graph, the architecture with the LLG design has the best utilization on LUT (36.72%), LUTRAM (3.13%) and FF (7.31%) components, while architecture LLB has the worst utilization of the three components 44.94% for LUT, 4.18% for LUTRAM and 9.40% for FF. Meanwhile, for BRAM utilization, LLG has the best utilization with 2.47% utilization and LLA has the worst utilization with 6.58% utilization. The percentage improvement between the least and best utilization of the four components is given in Table 4.

For ASIC implementation, the resource improvement of each design was estimated based on the number of cells, the total cell area and the size of the layout area. Table 5 shows the collected data for each architecture.

Based on Table 5, the architecture LLG has the smallest layout area which is 25.60 μm^2 and the architecture LLA has the largest layout area which is 43.96 μm^2 . This is a reduction of around 40% in layout area in the final design, compared to the initial design.

Table 4 Percentage improvements between the best and the worst for LUT, LU-TRAM, FF and BRAM components

Components	Worst (%)	Best (%)	Difference (%)
LUT	44.94	36.72	8.22
LUTRAM	4.18	3.13	1.05
FF	9.40	7.31	2.09
BRAM	6.58	2.47	4.11

Table 5 The number of cells, the total cell area and the size of the layout for all designs

Design	Cell count	Cell area (μm^2)	Layout area (μm^2)
LLA	983,660	33.26	43.96
LLB	939,028	32.02	41.99
LLC	929,612	31.40	41.09
LLD	833,534	27.78	36.48
LLE	707,334	22.46	31.47
LLF	649,237	19.96	28.30
LLG	619,439	18.53	25.60

4 Conclusion

The hardware design and implementation of BLE controller have been successfully performed using the CAL dataflow specification. As a result, seven different architectures were obtained. An optimized design of the BLE link layer controller is obtained by applying high-level optimization techniques—the state reduction techniques and actor merging techniques. However, different implementations show different performance parameters. For FPGA implementation, LLD architecture shows the most optimized power consumption and energy consumptions but not the resources. The most optimum resource utilization is shown with the LLG architecture. Meanwhile, for ASIC implementation, the best design with the most optimum power, energy, total number of cells and area is the LLG architecture.

The results also show that the GPP implementation has the highest power and energy consumption compared to FPGA and ASIC. Even though ASIC should theoretically show lower power and energy consumption compared to FPGA, the results show that the FPGA implementation achieved lower power than ASIC due to the technology used, i.e., 180 nm and 14 nm technology for ASIC and FPGA, respectively. In terms of speed, the GPP required more time to execute the test bench, while FPGA and ASIC have the same speed and total time to execute due to the same operating frequency applied for both implementations.


References

1. R. Heydon, *Bluetooth Low Energy: The Developers Handbook* (Prentice Hall, Upper Saddle River, 2012)
2. Bluetooth SIG, *Bluetooth Specification Version 4.2* (Bluetooth SIG Inc., 2014)
3. H. Amer, A.A.H. Ab-Rahman, I. Amer, C. Lucarz, M. Mattavelli, Methodology and technique to improve throughput of fpga-based cal dataflow programs: case study of the rvc mpeg-4 sp intra decoder, in *2011 IEEE Workshop on Signal Processing Systems (SiPS)* (2011), pp. 186–191
4. E. Bezati, M. Mattavelli, J.W. Janneck, High-level synthesis of dataflow programs for signal processing systems, in *2013 8th International Symposium on Image and Signal Processing and Analysis (ISPA)*, Trieste (2013), pp. 750–754

5. H. Yviquel, A. Lorence, K. Jerbi, G. Cocherel, A. Sanchez, M. Raulet, Orcc: multimedia development made easy, in *Proceedings of the 21st ACM International Conference on Multimedia* (2013), pp. 863–866
6. I.J. Chun, B.G. Kim, I.C. Park, Fully synthesizable bluetooth basedband. *ETRI J.* **25**(5) (2003)
7. Y. Yun, W. Danghui, Y. Ke, F. Zhihua, Design of link layer controller for high speed serial bus, in *The 2nd International Conference on Information Science and Engineering*, Hangzhou (2010), pp. 1997–2000
8. P. Wiecha, M. Cieplucha, P. Kloczko, W.A. Pleskacz, Architecture and design of a bluetooth low energy controller, in *2016 MIXDES—23rd International Conference Mixed Design of Integrated Circuits and Systems*, Lodz (2016), pp. 164–167
9. S. Raza, P. Misra, Z. He, T. Voigt, Building internet of things with bluetooth smart. *Ad Hoc Netw.* **57**, 19–31 (2016)
10. S. Das, S. Ganguly, S. Ghosh, R. Sarker, D. Sengupta, A bluetooth based sophisticated home automation system using smartphone, in *2016 International Conference on Intelligent Control Power and Instrumentation (ICICPI)*, Kolkata (2016), pp. 236–240
11. T. Zhang, J. Lu, F. Hu, Q. Hao, Bluetooth low energy for wearable sensor-based healthcare systems, in *2014 IEEE Healthcare Innovation Conference (HIC)*, Seattle, WA (2014), pp. 251–254
12. B.G.A. Kumar, K.C. Bhagyalakshmi, K. Lavanya, K.H. Gowranga, A blue-tooth low energy based beacon system for smart short range surveillance, in *2016 IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT)*, Bangalore (2016), pp. 1181–1184
13. N.M. Zabidi, A.A.H. Ab Rahman, VLSI design of a fast pipelined 8x8 discrete cosine transform. *Int. J. Electr. Comput. Eng.* **7**(3), 1430–1436 (2017)
14. A.A.H. Ab Rahman, I. Kamisian, A.Z. Shaameri, VLSI design and implementation of adaptive channel equalizer, in *2008 International Conference on Computer and Communication Engineering* (2008), pp. 1121–1124
15. M. Chavarras, F. Pescador, M.J. Garrido, A. Sanchez, C. Sanz, Design of multicore HEVC decoders using actor-based dataflow models and OpenMP. *IEEE Trans. Consum. Electron.* **62**(3), 325–333 (2016)
16. A.A.H. Ab Rahman, R. Thavot, M. Mattavelli, P. Faure, Hardware and software synthesis of image filters from cal dataflow specification, in *6th Conference on Ph.D. Research in Microelectronics & Electronics* (2010), pp. 1–4

Bandwidth and Gain Enhancement of Circular MPA Using Twin-Layer Stacked Antenna



Harshit Srivastava , Usha Tiwari, and Pallavie Tyagi

Abstract A microstrip patch antenna (MPA) has been nowadays employed in various types of applications. Apart from the various advantages that it has, MPA predominantly has two disadvantages; narrow bandwidth and low gain. As the patch's shape decides the type of the MPA, one such antenna is a circular microstrip patch antenna, having a circular-shaped patch. In this paper, a novel approach has been put forth in enhancing the gain and bandwidth of a circular antenna, simultaneously. This has been achieved using a novel antenna, i.e., "twin-layer stack antenna." This antenna has been made by deploying another layer of the substrate, having a dielectric value of 1.5, over the patch. Hence, with the help of another stacked layer of the substrate, we have endeavored to enhance the gain and bandwidth of the conventional circular MPA. Both the antennas have been designed to operate at a frequency of 7.6 GHz, which comes under the C-band of the electromagnetic spectrum. This particular band is generally used for wireless radio bands, WLAN, and Wi-Fi applications. Using the co-axial feeding technique and HFSS software, the results have been drawn. After simulation and further calculations, the bandwidth for circular MPA comes out to be 230 MHz, while for twin-layer stacked antenna, it is found out to be 240 MHz, i.e., an enhancement of 10 MHz. While the gain of the twin-stacked circular antenna comes out to be 2.6587 dB, whereas 0.9372 dB of gain for the circular MPA has been achieved, that means an enhancement of 1.7215 dB.

Keywords Gain · Bandwidth · Microstrip patch antenna · Stacked antenna · Circular MPA

H. Srivastava (✉)
Highlands, Noida, India

U. Tiwari
Sharda University, Greater Noida, India

P. Tyagi
ABES Engineering College, Ghaziabad, NCR, India

1 Introduction

The circular and rectangular microstrip patch antenna are the most commonly used MPAs. The reason behind this is that they provide linear/circular polarization, multiple frequency operations, feedline flexibility, and its flexibility to array configuration [1]. However, apart from these advantages, they have predominantly two significant disadvantages: low gain and narrow bandwidth, which somewhat affect the antenna's efficiency [2]. A novel approach has been put forward to increase the gain and bandwidth of a circular MPA/single-layer stacked antenna using a twin-layer stacked antenna. This twin-layer stacked antenna has two layers of the stacked dielectric substrate, placed over one another. The EM spectrum has been further diversified into many bands, which play an essential role in different communication approaches. In this paper, we have studied the C-band, which is used for wireless radio bands, WLAN, Wi-Fi, and ISM band applications. The resonant frequency is selected as 7.6 GHz that is most commonly used for wireless radio bands, WLAN, as well as for Wi-Fi applications. This band has come into the discussion since the late 1990s to till date as many people have given their contributions ranging from 4 to 8 GHz; still, there is much more to investigate into this band.

There are several primitive techniques in the market that enhances either the gain or the bandwidth of the antenna. The simplest approach for enhancing the bandwidth of an MPA can be done by either decreasing the dielectric constant of the substrate or to increase the height of the same. However, if the height of the substrate is increased further, then it no longer remains compact. In contrast, if the dielectric constant is decreased, then the antenna will not be able to resonate at higher frequencies, while the gain could be enhanced by using a low loss substrate. Hence, there are not any single technique which could enhance both the gain and bandwidth, simultaneously [2]. An endeavor has been put forth in the bandwidth and gain enhancement of the antenna at the same resonating frequency using a co-axial feeding technique, as discussed earlier. Although, apart from co-axial feeding, there are other techniques too which have been used extensively, such as aperture coupling, microstrip line, and proximity coupling [3]. The comparison has been drawn w.r.t. gain and bandwidth of both the antennas. It is quite notable that the return loss of the twin-layer stacked antenna has increased by 2 dB on comparing it with single-layer stacked/circular MPA. Designing of the two antennas has been covered in forwarding sections of the paper (Figs. 1 and 2).

2 Literature Review

In [4], the designing of the circular patch antenna is proposed using a stack-layering technique in which the operating frequency is 9 GHz using HFSS software. Here, a two-layer circular patch antenna has been proposed to enhance the antenna's gain and bandwidth using two substrates of different dielectric constants. It is observed

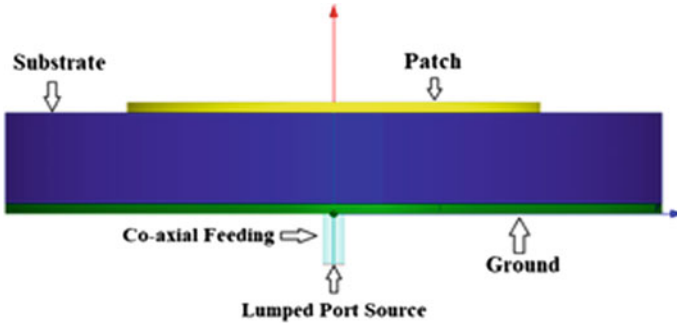


Fig. 1 Side view of the single-layer stacked antenna

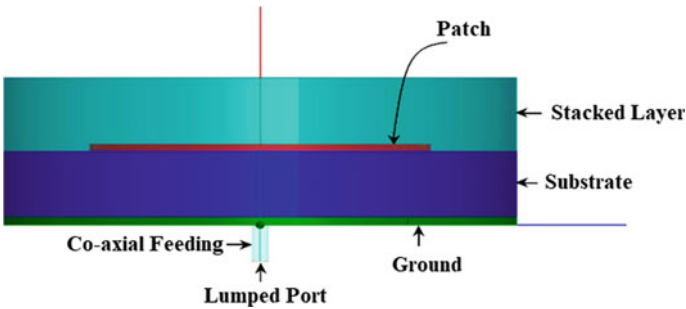


Fig. 2 Side view of twin-layer stacked antenna

that the antenna can achieve a good return loss for simple and double-layer stack antenna as -41.96 dB and -23.07 dB, respectively, with a gain enhancement of 4 dB. An endeavor has been put forth to enhance the gain of the antenna with the help of electromagnetic band gap (EBG) using finite difference time domain (FDTD) has been proposed in [5]. EBG helps in the recession of the propagation of surface waves at a particular frequency. The operating frequency is selected as 7.5 GHz, for which the gain is calculated for both rectangular MPA and stacked antenna. The obtained values for both the antennas are 6.30 dB by using the EBG structures which are responsible to get a high gain in the antenna. A similar study has also been put forward in [6] for enhancing the performance of a circular MPA via a cylindrical EBG substrate.

Not only in X-band, but the stacked antennas have also been studied for the S-band [7] as well in which 2.45 GHz is being selected as the operating frequency for the simulation through GENESYS software. The return loss and VSWR of the antenna at 2.45 GHz have been obtained as -19.98 and 1.22 dB, respectively, by taking the substrate's thickness at 0.5 mm. The impact of feeding [8] can also change some antenna parameters and is another factor for the enhancement of the antenna's bandwidth. In this proposed sample, a simple, economical, and most widely used

antenna has been designed, and the comparative study has been performed by varying the feeding techniques of the antenna. It has been observed that the proximity-coupled feeding technique is considered to be best as it provides up to 13% enhancement in bandwidth, whereas other feeding techniques result in 2–5% enhancement.

In [9], the bandwidth of the MPA was enhanced by replacing the conducting ground plane with a high impedance EBG layer. It was concluded that on increasing the radius of each EBG cell (circular), the bandwidth increased from 13.41 to 18.68%. A novel technique has been proposed in [10] where rectangular and fractal U-shaped slots have been introduced on the patch of the microstrip patch antenna. The antenna was designed to operate at 1.6 GHz and a gain enhancement of 18.6%, whereas a bandwidth enhancement of 100% was achieved. A circular polarized X-band microstrip array antenna has been proposed in [11] where the bandwidth was enhanced using the elliptical-ring-slot (ERS) technique. The proposed antenna incorporates circular shaped with truncation factor and ERS in the patch. In [12], a unique approach was proposed in which the circular MPA with an air gap between the two substrates and ground plane was done using the cavity model theory. The gain was enhanced by 2.39 dB, and the bandwidth was increased by 1.35% as compared to the conventional circular MPA. In [13], a polarization-reconfigurable rectangular MPA for enhancement of the bandwidth was proposed. The antenna incorporates stair-slots on the ground and two pin diodes for switching its polarization in linear, LHCP, and RHCP. This antenna also saw better switchable and radiation performance.

2.1 Design Parameters

The incumbent parameters for designing of the single and twin-layer stacked circular antenna using the co-axial feeding technique are as follows:

- Dielectric constant (ϵ_r) = 4.4
- Operating frequency (fr) = 7.5 GHz
- Thickness of substrate (H) = 10 mm
- Radius of patch (r) = 20 mm
- Loss tangent = 0.02

On varying the parameters, one can observe a shift in the operating frequency [14].

3 Single-Layer Stacked Circular Antenna

See Fig. 3

Figure 4 shows the return loss of the single-layer stacked circular antenna. It can be inferred from the figure that the return loss for the antenna is -22.6745 dB.

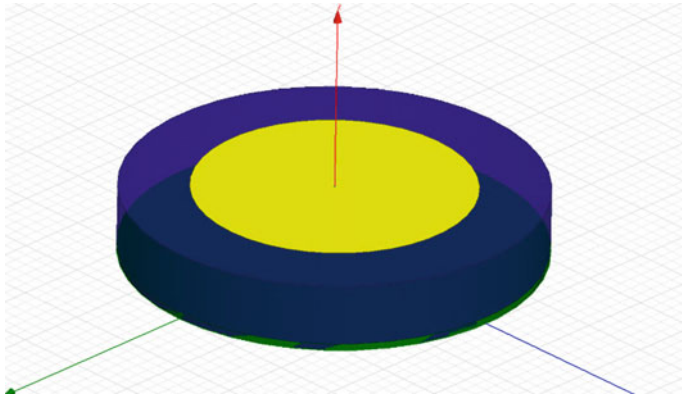


Fig. 3 Proposed single-layer stacked circular antenna using co-axial feeding

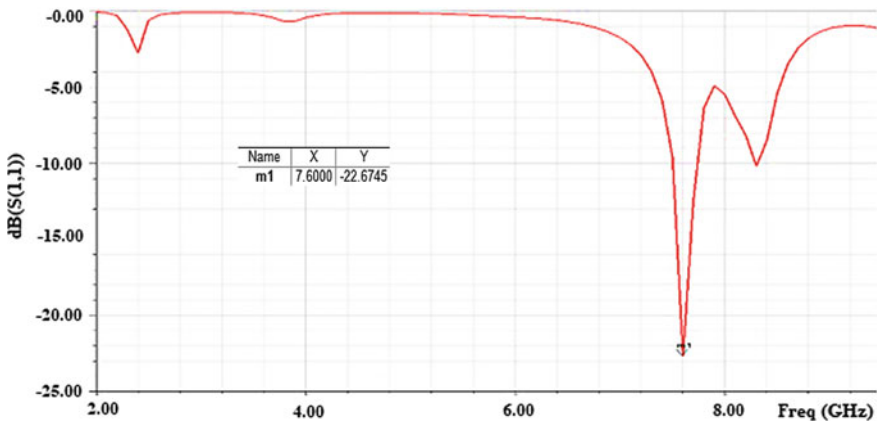


Fig. 4 Return loss for the single-layer stacked circular antenna

Figure 5 represents the bandwidth of the proposed antenna. It can be noticed that the bandwidth of the single-layer stacked circular antenna comes out to be 230 MHz, which is a narrow bandwidth.

Figures 6 and 7 show the gain versus theta and radiation plot of the single-layer stacked circular antenna, respectively, from which it is noted that the gain comes out to be 0.9372 dB, for $\theta = 0^\circ$.

4 Twin-Layer Stacked Circular Antenna

Figure 8 shows the diagrammatic representation of the proposed twin-layer stacked circular antenna in HFSS software. For this, two substrates have been used one is of

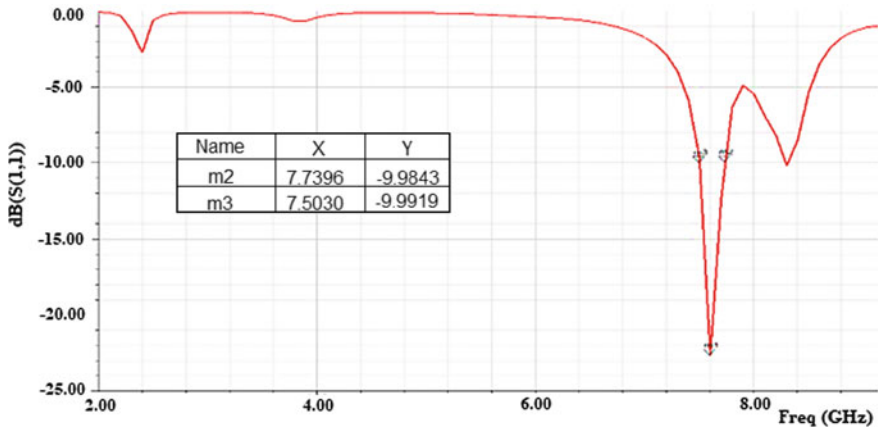


Fig. 5 Bandwidth of the single-layer stacked circular antenna

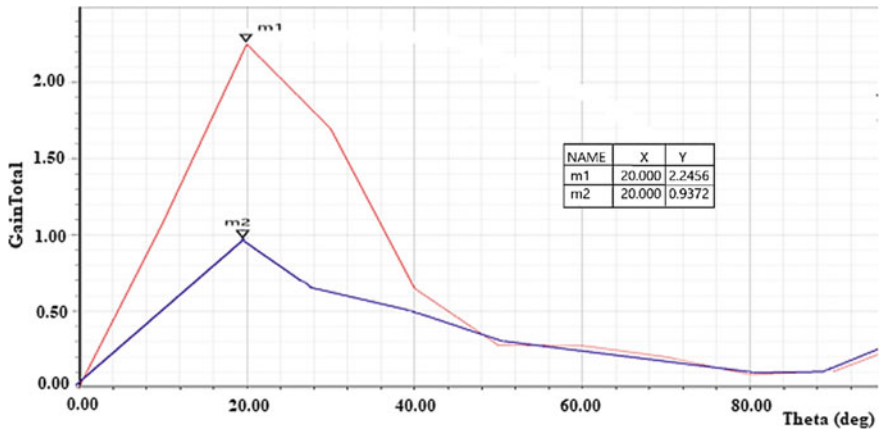


Fig. 6 Gain versus theta plot for single-layer stacked circular antenna

dielectric constant 4.4, and the other one is having $\epsilon_r = 1.5$. From Fig. 9, it can be seen that the return loss for the twin-layer stacked circular antenna is shown which is found out to be -20.0456 dB.

From Fig. 10, the bandwidth for the twin-layer stacked circular antenna can be obtained, which comes out to be 240 MHz.

At last, the gain for the twin-layer stacked circular antenna can be inferred from Figs. 11 and 12, which is observed as 2.6581 dB, for $\theta = 0^\circ$.

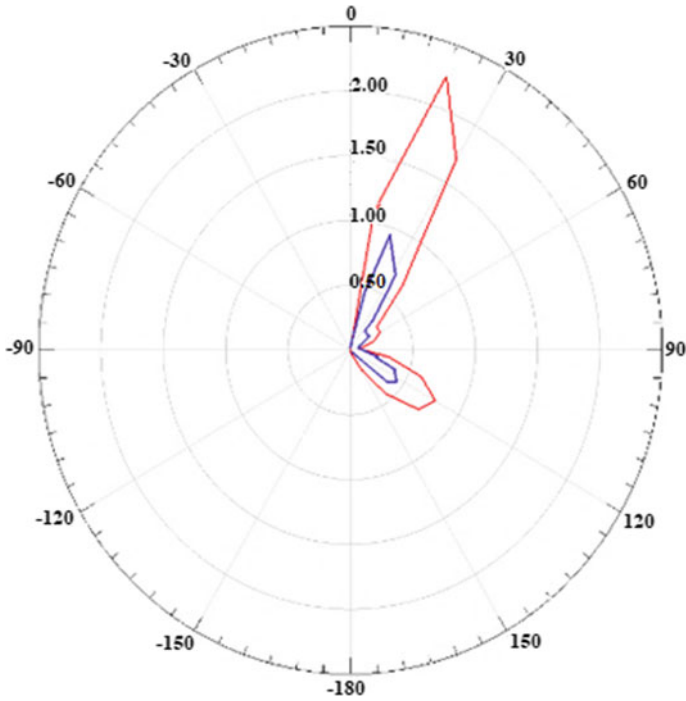


Fig. 7 Radiation pattern for single-layer stacked circular antenna

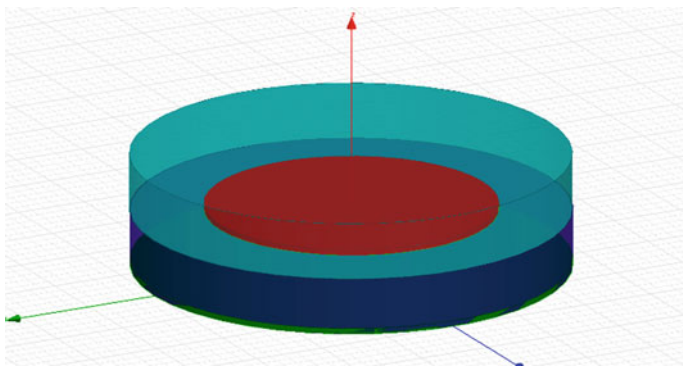


Fig. 8 Proposed twin-layer stacked circular antenna using co-axial feeding

5 Discussion

The proposed novel approach has been proven to enhance both the gain as well as the bandwidth of the single-layer circular antenna simultaneously. It is quite known that both bandwidth and resonating frequency are inversely proportional to the value of

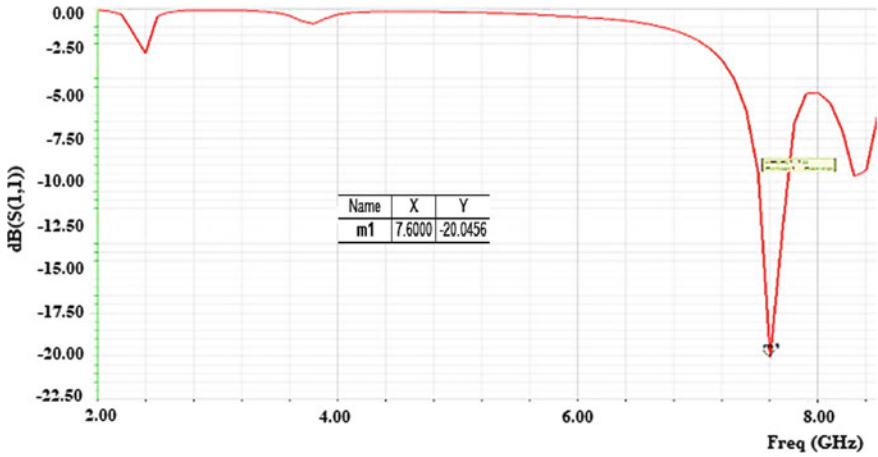


Fig. 9 Return loss for twin-layer stacked circular antenna

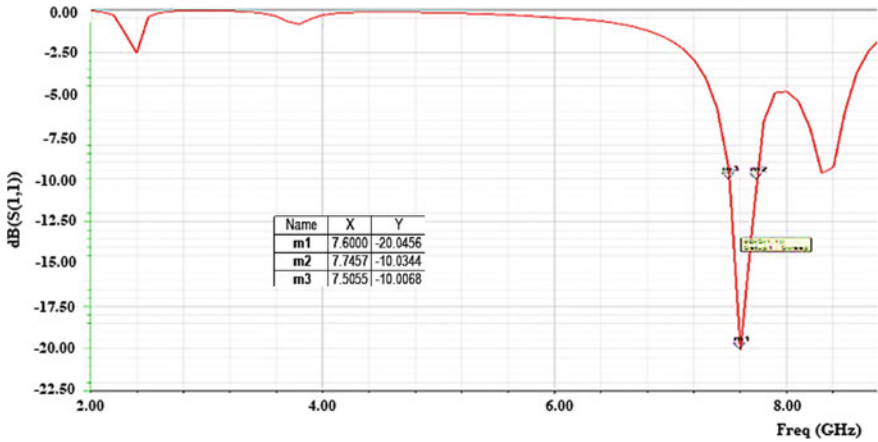


Fig. 10 Bandwidth of twin-layer stacked circular antenna

dielectric constant, making the proposed antenna work at a frequency of 7.6 GHz, a dielectric material having 4.4 as the constant value was chosen. Moreover, it should be noted that all the results have not been verified with the fabricated antenna. The conclusion that has been drawn is solely based on simulation results; however, the fabrication part could be done in the future to validate the results. As everything has its own pros and cons, this novel approach has a disadvantage too that came into the picture while seeing the results for return loss. It was noted that the return loss of approximately 2 dB was found more in twin-layer stacked antenna as compared to the single-layer circular antenna (Table 1).

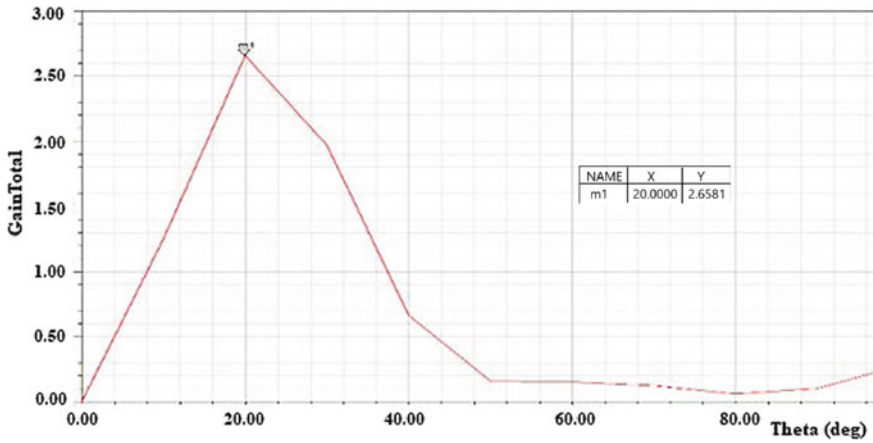


Fig. 11 Gain versus theta plot for twin-layer stacked circular antenna

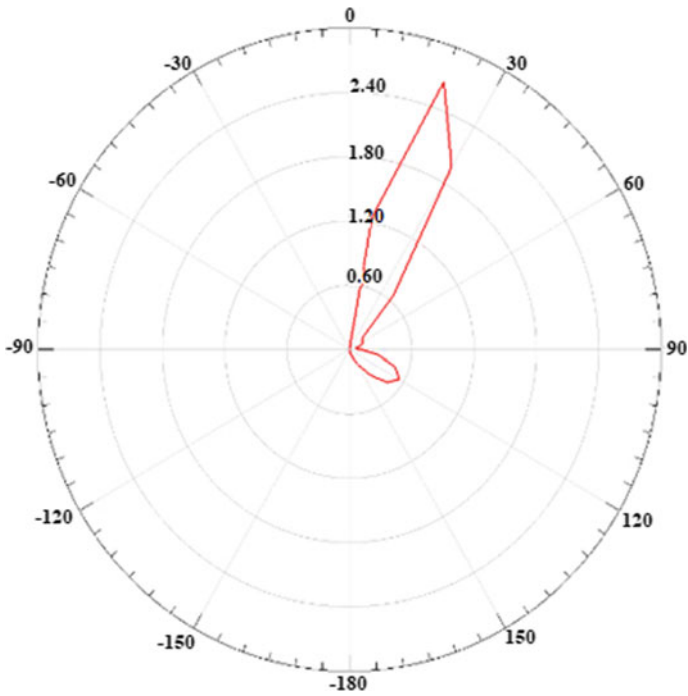


Fig. 12 Radiation pattern for twin-layer stacked circular antenna

Table 1 Analogy of various parameters of performance

Performance parameters	Single-layer circular antenna	Twin-layer stacked antenna
Return loss (dB)	-22.674	-20.045
Bandwidth (MHz)	230	240
Gain (dB)	0.9372	2.6587

6 Conclusion

Design, simulation, and comparison of single-layer stacked circular antenna and twin-layer stacked circular antenna using co-axial feeding method for an operating frequency of 7.6 GHz are performed in this paper. This is achieved with the help of HFSS software. The return loss, bandwidth, and gain of the single-layer stacked antenna are observed as -22.674 dB, 230 MHz, and 0.9372 dB, respectively, whereas, for the twin-layer stacked antenna, the return loss, bandwidth, and gain are obtained as -20.0456 dB, 240 MHz, and 2.6581 dB, respectively. It can be seen that the single-layer stacked antenna is dominating over twin-layer stacked circular antenna in terms of return loss, but as we check for the other two parameters, it is visible that the bandwidth of the antenna has improved for twin-layer circular antenna from 23 to 240 MHz and the gain has increased from 0.9372 to 2.674 dB. Hence, bandwidth and gain enhancement have been achieved in this paper. The proposed antennas can also be used in the wireless radio band, WLAN, as well as for Wi-Fi applications.

References

1. H. Srivastava, U. Tiwari, Design, simulation and analysis of rectangular and circular microstrip patch antenna for wireless applications. *Int. J. Recent Technol. Eng. (IJRTE)* **8**(4), 2277–3878 (2019)
2. H. Srivastava, A. Singh, A. Rajeev et al., Bandwidth and gain enhancement of rectangular microstrip patch antenna (RMPA) using slotted array technique. *Wireless Pers. Commun.* (2020). <https://doi.org/10.1007/s11277-020-07388-x>
3. H. Srivastava, A. Singh, A. Rajeev, U. Tiwari, Comparison of different microstrip patch antennas with proposed RMPA for wireless applications, in 2019 International Conference on Power Electronics, Control and Automation (ICPECA), New Delhi, India, pp. 1–4 (2019). <https://doi.org/10.1109/ICPECA47973.2019.8975599>
4. A. Aggarwal, S. Khurana, Wzir, Design and simulation of dual stacked antenna used in wireless communication. *Int. J. Eng. Comput. Sci.* **3**(6), 6491–6494 (2014)
5. N. Rao, D. Kumar, Investigation of a Microstrip Patch Antenna with EBG Structures Using FDTD. V978-1-4244-9477-4/11/\$26.00 ©2011 IEEE
6. H. Boutayeb, T.A. Denidni, Gain enhancement of a microstrip patch antenna using a cylindrical electromagnetic crystal substrate. *IEEE Trans. Antennas Propag.* **55**(11), 3140–3145 (2007). <https://doi.org/10.1109/TAP.2007.908818>
7. P.L. Zade, S.S. Khade, N.K. Choudhary, Modeling and designing of circular microstrip antenna for wireless communication, in *Proceedings of the 2009 Second International Conference on Emerging Trends in Engineering & Technology*, pp. 1189–1194, December 2009

8. S. Singh, N. Agarwal, N. Nitin, A.K. Jaiswal, Design consideration of microstrip patch antenna. *Int. J. Electron. Comput. Sci. Eng.*, ISSN 2277-1956
9. R. Gupta, M. Kumar, Bandwidth enhancement of microstrip patch antennas by implementing circular unit cells in circular pattern, in *2013 5th International Conference and Computational Intelligence and Communication Networks*, Mathura, pp. 10–13 (2013). <https://doi.org/10.1109/CICN.2013.11>
10. P.P. Chandran, S. Viswasom, Gain and bandwidth optimization of a novel microstrip patch antenna, in *2014 Fourth International Conference on Advances in Computing and Communications*, Cochin, pp. 315–318 (2014). <https://doi.org/10.1109/ICACC.2014.80>
11. F. Kurniawan, J.T.S. Sumantyo, A. Bintoro, D.A. Purnamasari, Bandwidth enhancement of circular polarized X-band microstrip array antenna using ERS, in *2017 IEEE Conference on Antenna Measurements & Applications (CAMA)*, Tsukuba, pp. 228–231 (2017). <https://doi.org/10.1109/CAMA.2017.8273409>
12. T. Divakar, D.C. Panda, Gain and bandwidth enhancement of a circular microstrip patch antenna with an air dielectric between two substrates, in *2013 IEEE Applied Electromagnetics Conference (AEMC)*, Bhubaneswar, pp. 1–2 (2013). <https://doi.org/10.1109/AEMC.2013.7045098>
13. Z. Yang, H. Yang, J. Hong, Y. Li, Bandwidth enhancement of a polarization-reconfigurable patch antenna with stair-slots on the ground. *IEEE Antennas Wirel. Propag. Lett.* **13**, 579–582 (2014). <https://doi.org/10.1109/LAWP.2014.2312971>
14. Co.A. Balanis, *Antenna Theory, Analysis and Design*, 4th edn. (Wiley, 2015)

Rooftop Antenna for Vehicular Application



Muskan Bhatia, Mansi Dimri, and Brajlata Chauhan

Abstract This paper presents design and analysis for microstrip planar antenna and conformal antenna, which operates in X & ku band for SOTM application. The design has been made in such a way so that surface current can be reduced, and it is designed to be fed by a dual port. A comparison between results observed for different materials used in microstrip patch antenna and conformal antenna has been tabulated. Comparison between different materials showed duroid 6006™ to be desired material having 6.15 dielectric constant and thickness of 0.7 mm. This design can be seen to be operating in three bands—X, Ku, and K band.

Keywords Antenna · SOTM · Microstrip patch · Conformal antenna · HFSS

1 Introduction

SATCOM on the Move (SOTM), or satellite correspondences moving, is an expression utilized with regards to portable satellite innovation, explicitly identifying with military ground vehicles, maritime, and airborne stages. The essential rule behind SATCOM on the Move is that a vehicle outfitted with a satellite antenna can build up a connection with a satellite and keep up that connection while the vehicle is moving with its well-defined system with two of its key components—antenna and modem [1]. SOTM systems can be of various types like satellite—instant worldwide coverage satellite Communications (SATCOM), commercial maritime and land mobile fleet management, Wi-Fi/cellular distribution for mass passenger transportation platforms, and military mobile satellite communications. This type of system also faces many types of challenges like the line of sight, cost, and complex manufacturing [2].

To cover up this challenge, most vehicle makers utilize the rooftop to put an antenna. This has an undeniable explanation, since the top of a vehicle is high over the ground and unhindered providing better reception, where conformed antennas come

M. Bhatia · M. Dimri · B. Chauhan (✉)
DIT University, Dehradun, Uttarakhand, India

in handy. Vehicle-to-car communication and car-to-infrastructure communication are right now one of the most mainstream fields of investigates for effective vehicle data frameworks [3] “Conformal antennas” came into existence in the 1980s. A conformal is an antenna that adjusts to some recommended shape like a cylinder, cone, sphere, human body, and train [4, 5].

The objective is to build up the conformal antenna for SOTM application which can be conformed to any vehicular surface such that it does not create a hindrance. Planar microstrip antenna is best suited for this job. MSA is simple design with a dielectric layer sandwiched between two layers: ground and a radiating patch lying on the top. Microstrip antenna can be fed with different types of feeding techniques: microstrip feeding, coaxial probe, aperture coupled, and proximity coupled.

Moreover, a middle layer of the antenna hampers the performance of the antenna as the thickness of the dielectric layer, but with a lower dielectric constant is inversely proportional to the performance. MSA supports multiple bands of frequency, is light in weight, supports dual-polarization, low manufacturing cost, robust nature,

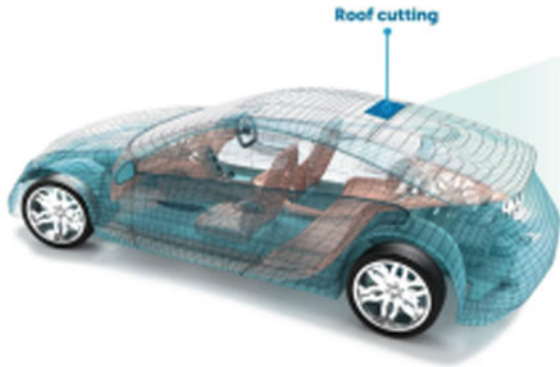


Fig. 1 Vehicle designed with rooftop antenna [6]

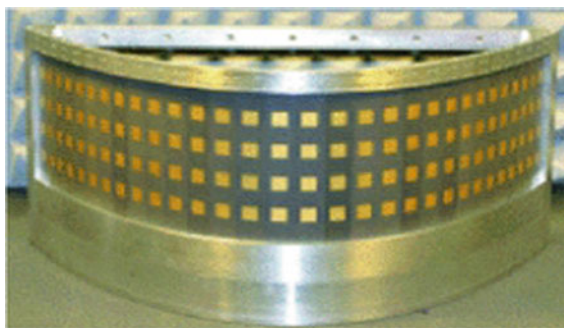


Fig. 2 Array of antennas conformed on surface [7]

operation at microwave frequencies, and may more [8], but has narrow bandwidth [9].

The motive behind using a conformal antenna in the vehicular application is majorly to make it less visible and perform its task efficiently; this can be fulfilled by choosing a microstrip patch antenna to conform to a prescribed shape like a cylinder (common), cone, sphere, etc. All the results have been simulated and tabulated. All the results have been simulated in HFSS and tabulated with graphs.

2 Vehicular Rooftop Antenna

Conformal antenna can be designed for many applications like it can be designed for V2X (vehicle to everything). Earlier, a new structure of antenna was developed with Yagi-Uda array concept and the microstrip radiator technique for the mobile satellite (MSAT) system as a low-profile, low-cost antenna [10] (Table 1).

The Low-profile conformable antenna on a regular surface, viz. cylindrical surface, is easily achievable by conforming microstrip patch antenna on the surface. Design procedure involves parameters related to cylindrical microstrip elements to realize the desired resonant frequency, input impedance, radiation patterns, etc. [11] The antenna can also adopt the structure of an L-shaped monopole loaded by a shorted pin to realize characteristics of low-profile and vertical polarization [12] or it can be shark-fin or a conformal antenna which can be used in different types of systems like GPS [13], LTE, cellular frequency, etc. [11]. The conformal antenna may show different bandwidths for different angles and radii with the same microstrip antenna, whereas further improvement can be done which may result in better bandwidth [14].

Table 1 Review analysis

S. No.	Proposed by	Material	ϵ_r	Freq (GHz)	Gain (dB)	B.W
1	Sameer Dev Gupta	RT/duroid 5870	2.2	10	6.9	5.8%
2	Zong-Quan Liu	Air	1	820 MHz	1.5	28 MHz
3	Prateek Chopra	RT/duroid 5880	2.2	9.6	2.75	148 dB
4	Martin Wooseop Lee	Copper	1	0.78 2.4 3.5 5.9	3.23 5.8 5.63 4.93	–
5	John Huang	Air	2.5	1.654	10	–

3 Planar Microstrip Patch Antenna

Most commonly available and used dielectric materials for a substrate with ability to conform are used [15]. In this work, a microstrip patch antenna with patch length 9 mm and width as 8 mm and parallel bars length 6.5 mm and width as 1.2 mm which then fed with dual-port is simulated at 9.2 GHz. For the designing of the antenna, ANSYS HFSS 13 has been used. All the results have been tabulated.

In Table 2, all the dielectric materials have been listed out with their values of relative permittivity (Figs. 3, 4, 5, 6 and 7).

The above results depict the return loss graph for different materials used as substrates with their corresponding gain plots (dB). For the microstrip patch, return loss is maximum at -40.43 dB for duroid 6002TM followed by duroid 5870, duroid 5880, and duroid 6006 with a maximum bandwidth of 4.4 GHz. Antenna with Rogers RT/duroid shows operation at two frequencies in X and Ku band with frequencies, 6.3 GHz and 13.3 GHz, respectively. The gain for the above design has been found around -5.603 dB.

Table 2 Frequency and S parameter for different materials

Material	ϵ_r	Frequency (GHz)	dB (S (1,1))
RT duroid 5870 TM	2.33	9.8 9.9	-38.2465 -38.3104
RT duroid 5880 TM	2.2	10.1	-38.7972
RT duroid 6002 TM	2.94	8.9	-40.4327
RT duroid 6006 TM	6.15	6.3 13.9	-38.7090 -25.9771

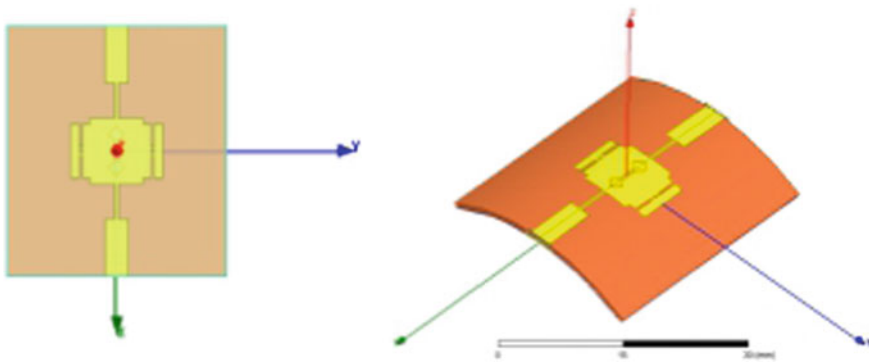


Fig. 3 Microstrip planar patch antenna and conformal microstrip patch antenna design

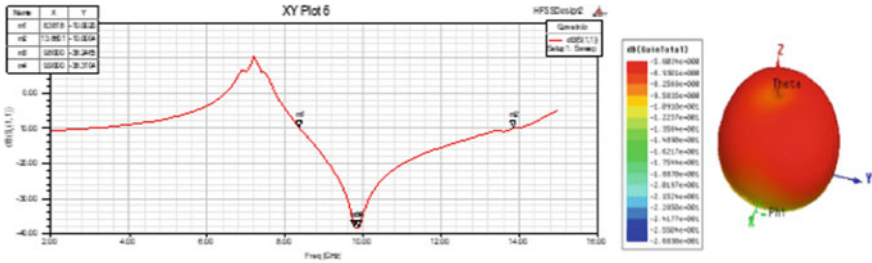


Fig. 4 dB (S (1,1)) and gain—Rogers RT/duroid 5870™

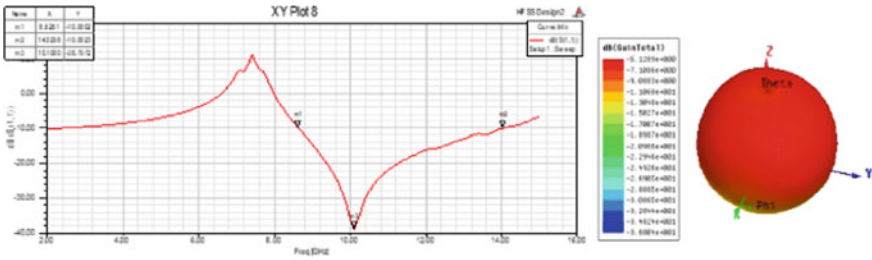


Fig. 5 dB (S (1,1)) and gain—Rogers RT/duroid 5880™

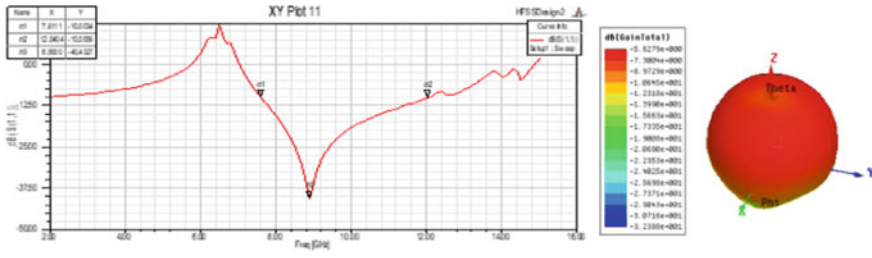


Fig. 6 dB (S (1,1)) and gain—Rogers RT/duroid 6002™

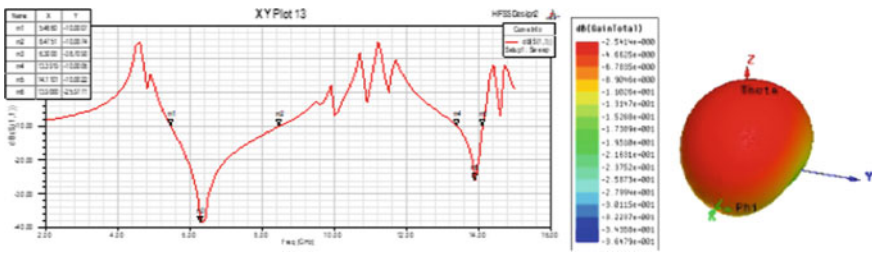


Fig.7 dB (S (1,1)) and gain—Rogers RT/duroid 6006™

4 Conformal Microstrip Patch Antenna

The proposed planar design has been conformed on a cylindrical surface with radius being 40 mm, and different changes in parameter has been observed for various flexible materials available which can be used for conforming. The proposed antenna has been simulated for X band and Ku band (Figs. 8, 9, 10, 11, and 12).

From the above graph, all the results are given in Table 3. Rogers RT/duroid 6006™ was found to have the maximum gain value of 5.517 dB and return loss of -12.591 and -15.2713 dB depicting its narrow band operation for frequencies 18.3 GHz and 20.1 GHz. All conformal designs are seen to be working in multiple frequencies in different bands. The observed angular width is 33.1223° ; the design can be steered up to 30° more.

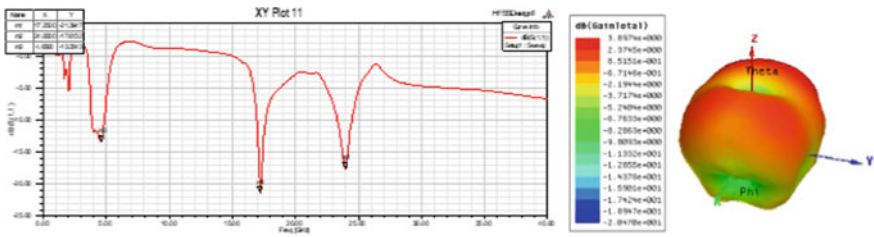


Fig. 8 dB (S (1,1)) and gain—Rogers RT/duroid 5870™

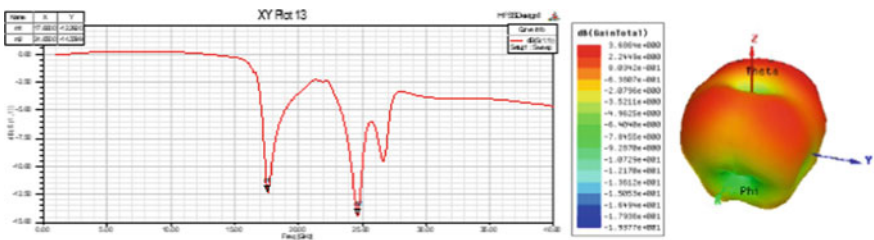


Fig. 9 dB (S (1,1)) and gain—Rogers RT/duroid 5880™

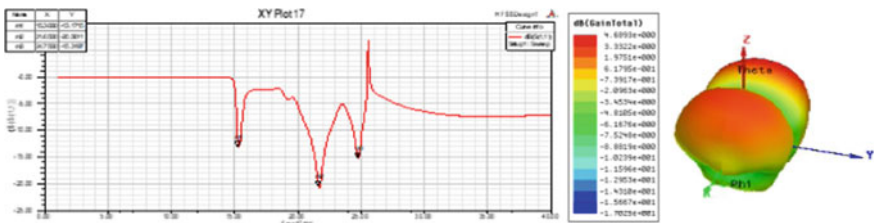


Fig. 10 dB (S (1,1)) and gain—Rogers RT/duroid 6002™

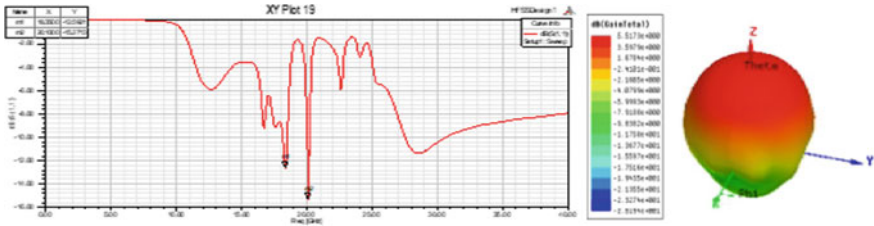


Fig. 11 dB (S (1,1)) and gain—Rogers RT/duroid 6006™

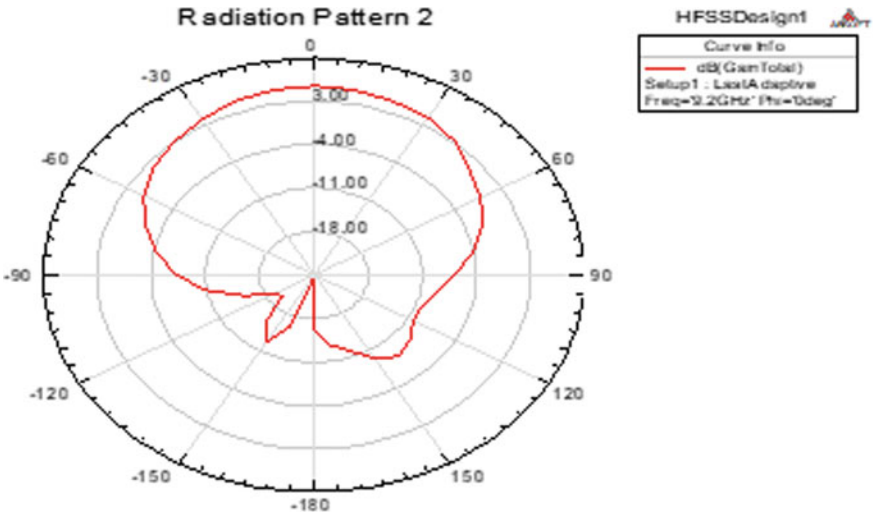


Fig. 12 Radiation pattern

Table 3 Dielectric constant, Gain, frequency, and S parameter for different materials

Material	ϵ_r	Gain (dB total)	Frequency (GHz)	dB (S (1,1))
RT duroid 5870™	2.33	3.8974	4.65	-13.3543
			17.25	-21.3947
			24	-17.6132
RT duroid 5880™	2.2	3.6864	17.6	-12.2920
			24.65	-14.3399
RT duroid 6002™	2.94	4.6893	15.3	-13.1715
			21.65	-20.5611
			24.75	-15.3197
RT duroid 6006™	6.15	5.5173	18.35	-12.5921
			20.1	-15.2713

5 Conclusion

The origination and simulation of microstrip planar and conformal antenna have been successfully designed via. the ANSYS high-frequency structure simulator (HFSS) operating in X band.

From the above-simulated results, for microstrip patch antenna, Rogers RT/duroid 6002™ gives better results for S_{11} , i.e., -40.4327 , for 10.1 GHz frequency achieving bandwidth of 4.4 GHz, working in X band, whereas gain has been degraded.

For conformal microstrip patch antenna from one of the four materials selected, antenna operates better for Rogers RT/duroid 6006™ with gain of 5.17 dB with a return loss of -12.5921 dB and -15.2713 dB at frequencies 18.35 GHz and 20.1 GHz, respectively, and with the angular width of 33.1223° .

From the above comparison, it is clearly observed that the conformal antenna with a dielectric material as duroid 6006™ is best suited to work in K band for higher order work—satellite communication (SATCOM). As, the viewpoint work, we may stretch out our research to the optimization of design and selection of material for better bandwidth and gain.

References

1. L. Gonzalez, M. Tracey, W. Lui, R. Greel, V. Aghinii, How to evaluate SOTM antennas, modems, and architectures, pp. 1–6 (2012). <https://doi.org/10.1109/MILCOM.2012.6415589>
2. M. Gachev, V. Boyanov, S. Kamenopolsky, V. Peshlov, B. Marinov, P. Dankov, On-the-move antenna systems for broad-band satellite communications, in *The 8th European Conference on Antennas and Propagation (EuCAP 2014)* (2014). <https://doi.org/10.1109/eucap.2014.6902312>
3. N. Koch, Antennas for Automobiles. New Advances in Vehicular Technology and Automotive Engineering (2012)
4. B. Chauhan, M. Bhatia, A. Chakraborty, P.S. Sharma, A. Sharma, Cylindrical conformal antenna arrays theory for military aircraft antenna, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 77–82 (2020). <https://doi.org/10.1109/GUCON48875.2020.9231232>
5. B. Chauhan, et al., Four element conformal array antenna for X-band and Ku-band applications. *IJCSIS Int. J. Comput. Sci. Inf. Secur.* **14**(5) (2016)
6. <https://www.fabbaloo.com/blog/2020/8/5/conformal-antennas-and-3d-printing>
7. P. Knott, T. Bertuch, H. Wilden, O. Peters, A.R. Brenner, I. Walterscheid, SAR experiments using a conformal antenna array radar demonstrator. *Int. J. Antennas Propag.* **2012**, 1–7 (2012). <https://doi.org/10.1155/2012/142542>
8. C.A. Balanis, *Antenna Theory: Analysis Design*, 3rd edn. (Wiley, New York, 1997). ISBN 0-471-66782-X 2005
9. A. Negi, B. Chauhan, V. Sandip, A microstrip patch antenna [pi] and T-shaped slot for X-band application. *IUP J. Telecommun.* **8**(2), 57–61 (2016)
10. J. Huang, A.C. Densmore, Microstrip Yagi array antenna for mobile satellite vehicle application. *IEEE Trans. Antennas Propag.* **39**(7), 1024–1030 (1991). <https://doi.org/10.1109/8.86924>
11. S.D. Gupta, et al., Design and performance analysis of cylindrical microstrip antenna and array using conformal mapping technique. *Int. J. Commun. Eng. Appl. IJCEA* **02**(03), 166–180 (2011)

12. C.M. del Redondo Gonzalez, Analysis of conformal antennas for avionics applications. Ph.D. thesis, Hardware for Wireless Communications, Chalmers University of Technology (2007)
13. E. Ghafari, A. Fuchs, D. Eblenkamp, D.N. Aloï, A vehicular rooftop, shark-fin, multiband antenna for the GPS/LTE/cellular/DSRC systems, in *2014 IEEE—APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)* (2014). <https://doi.org/10.1109/apwc.2014.6905546>
14. P. Chopra, M. Bhandari, K.S. Beenamole, S. Saxena, Design of an X- band conformal antenna using microstrip patches, in *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*. <https://doi.org/10.1109/spin.2015.7095393>
15. E. Abishek, A. Veera, K. Kumar Pc, A. Stephen, A. Raaza, Study and analysis of conformal antennas for vehicular communication applications. *ARPN J. Eng. Appl. Sci.* **12** (2017)

Design of Compact Quad Band Monopole Antenna Using Open-Ended and Meandered Slot



Raj Kumar and Yedukondalu Kamatham

Abstract A quad band with dual T-shaped stub with meander loaded slot antenna is designed to cover S and C band applications. The antenna generates quad resonant modes by etching symmetrically two open-ended rectangular slots and meandered slot. The radiating element with microstrip line with meandered-shaped slot and open-ended slot covers a frequency band 2.15–2.55 GHz, 2.75–3.7 GHz, 5.4–5.6 GHz, and 6.4–6.6 GHz at around 2.45 GHz, 3.1 GHz, 5.5 GHz, and 6.5 GHz, respectively. Also, the proposed design occupies a total area of $20 \times 18 \text{ mm}^2$, which is compact.

Keywords Quad band · Compact · Open-slot · Meandered-shaped slot

1 Introduction

Recent years, rapid growth in the evolution of wireless technology, demanding various application standards in modern wireless communication systems. To realize effective communication with various bands, multiband antenna design has accomplished a promising technology for wireless applications. Especially, a design of compact multiband antennas is playing a significant role [1–3] for various wireless applications such as Bluetooth (2.4–2.48 GHz), WiMAX (3.3–3.7 GHz), Unlicensed National Information Infrastructure (U-NII) WLAN (5.15–5.35 GHz), U-NII Wi-Fi (5.47–5.725 GHz), and U-NII WLAN (5.725–5.875 GHz) operations. A printed patch antenna is one of the important candidates for multiband operations [4]. However, the design of an antenna that simultaneously supports different

R. Kumar (✉)

Department of ECE, Shri Jagdishprasad Jhabarmal Tibrewala University, Vidyanagari, Rajasthan 333001, India

Department of ECE, Swami Vivekananda Institute of Technology, Secunderabad, Telangana, India

Y. Kamatham

Department of ECE, CVR College of Engineering, Hyderabad, Telangana, India

service bands with a single radiating element is a challenging task. Various multi-band monopole antennas have been discussed [5–10]. A trapezoidal ground plane and asymmetrical strip monopole antenna [5] and defected ground structure and inverted L-shaped [6, 7] have been designed to meet tri-band operation requirement. Some other tri-band antennas realized by cutting slots on radiating element [8–10]. However, the designs discussed above are only for WiMAX/ WLAN and WLAN/ ITU-T [10] applications and some of them are relatively more structural complexity and larger dimensions. Further, a less work has been reported in the literature for more than triple band applications. Recently, some quad band and five band antennas have been reported [11–14]. But the designs occupy larger dimensions than the proposed design presented in this paper.

In this paper, a small ($20 \times 18 \text{ mm}^2$) quad band antenna to cover multiservice applications is proposed. Owing to compact, the design can be a simple geometry for practical engineering applications. The antenna lower operating band is excited from the basic radiating element, and by incorporating two open-ended rectangular slots symmetrically on rectangular radiating patch additional symmetrical stub resonator is formed to generate resonance frequency of 5.7 GHz. Additionally, a meandered slot introduces a notch at 2.65 GHz and resonance at 6.5 GHz. The antenna is designed, and simulations are performed with the high-performance electromagnetic (EM) solver. Prototype of the design is fabricated and measured for comparison with simulation results.

2 Antenna Design

Figure 1 shows the proposed configuration of quad band antenna. The detailed dimensions shown in Fig. 1 are as follows (all dimensions are in mm): $W = 18$, $L = 20$, $W_1 = 16$, $W_2 = 7.5$, $W_3 = 6.0$, $W_4 = 4$, $W_f = 2$, $L_1 = 3$, and $L_2 = 7$. The structure is designed on FR-4 substrate with permittivity of 4.4 using finite element-based (FEM) commercial electromagnetic solver high-frequency structural simulator. The proposed compact monopole antenna is the ability to produce quad band resonance characteristics by presenting open-ended slots on the patch. The two open-ended slots named as rectangular and meandered slots are etched at different positions on a patch antenna. By etching, these slots generate three additional resonance frequencies, together with original monopole resonance.

Further, the design steps to achieve quad resonant modes of the antenna are shown in Fig. 2. The radiator with rectangular patch is designated as Antenna #1, which is operating at 2.8 GHz. An open-ended slot is etched on antenna #1 to create another resonant mode. This open-ended slot antenna is considered as Antenna #2 and operating at two different resonances with slightly shift in resonance of antenna #1. This is due to the slot cut on the patch, alters the current distribution which slightly shifts existing resonance and generated another resonance. In the next step, a meandered slot is etched on Antenna #2 to form Antenna #3 (Fig. 2c). After the meander slot, the antenna modified as a triple T-shaped stub resonator with meander arm. The middle

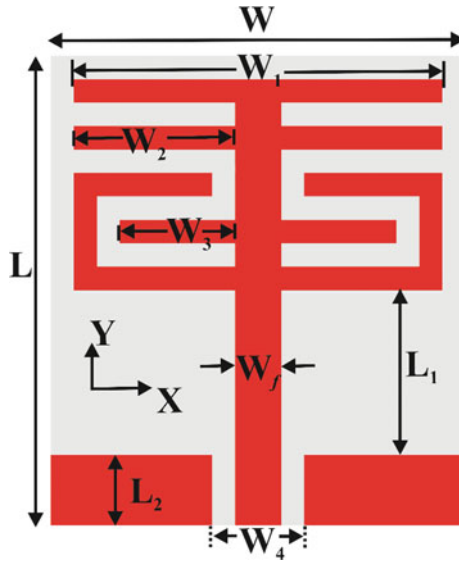


Fig. 1 Open-ended and meandered slot quad band antenna geometry

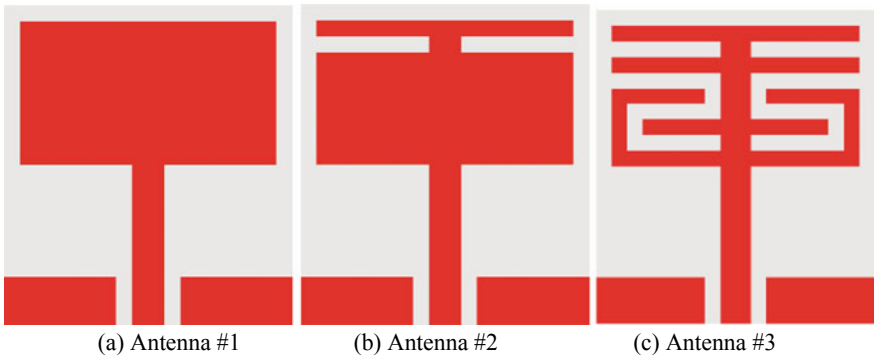


Fig. 2 Antenna design procedure in stepwise for quad band resonance

T-shaped resonator dimensions are the same as the top T-shaped resonator dimensions. So, there is no effective variation on resonance due to the middle T-shaped resonator. But, the third T-shaped resonator and meander arm generate an additional two resonance frequencies with slightly shifts to previous two resonant modes. The remaining two resonance frequencies are at 3.1 GHz and 6.5 GHz.

3 Results and Discussion

The simulated S_{11} of the antennas from Antenna #1 to Antenna #3 is shown in Fig. 3. Antenna #1 is operating on resonance frequency at 2.8 GHz over a frequency range from 2.3 to 3.3 GHz. While Antenna #2 is operating at 2.7 GHz (2.2–3.3 GHz) and 5.7 GHz (5.65–5.75 GHz) band. Whereas Antenna #3 is operating at 2.45 GHz (2.15–2.55 GHz) Wi-Fi/Bluetooth band, 3.1 GHz (2.75–3.7 GHz) WiMAX, 5.5 GHz (5.4–5.6) GHz WLAN band, and 6.5 GHz (6.4–6.6 GHz) mobile/ fixed satellite bands.

Further, the simulated current distribution of antenna #3 at 2.45 GHz, 3.1 GHz, 5.5 GHz, and 6.5 GHz is presented in Fig. 4. At 3.1 GHz is the distribution which is concentrated on the feed line and bottom part of the stubs. This leads to the elimination of certain frequencies creates another resonance frequency. Whereas, at 5.5 GHz, it is concentrated on upper part of the T-shaped stubs and middle part of the antenna. This indicates that the three resonance frequencies of Antenna #3 are

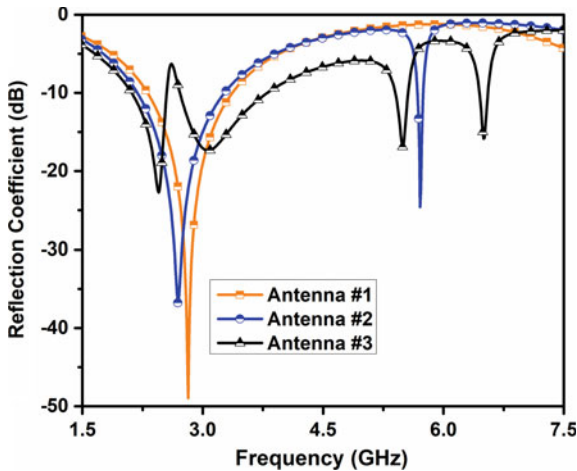


Fig. 3 Simulated reflection coefficient comparison

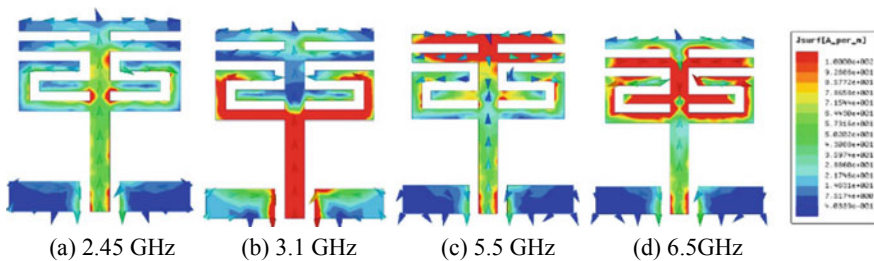


Fig. 4 Various resonant modes of surface current distribution

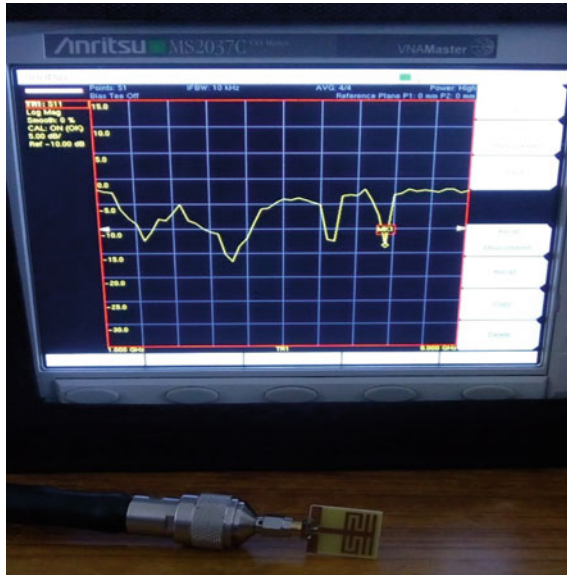


Fig. 5 Measured reflection coefficient of Antenna #3

due to feed and side arms of the antenna, upper T-shaped arms, and bottom side arms together with middle arms.

In order to verify the above-proposed quad band design strategy, the prototype is fabricated and measured the reflection coefficient. The measured reflection coefficient against a frequency is shown in Fig. 5. With negligible deviation, an agreement is attained to the simulation. The measured reflection coefficient of four resonance bands below -10 dB is observed at impedance bandwidth of 0.4 GHz, 1.05 GHz, 0.2 GHz, and 0.25 GHz at 2.15–2.55 GHz, 2.75–3.7 GHz, 5.4–5.6 GHz, and 6.4–6.6 GHz bands, respectively. The measured reflection coefficient of four resonance bands below -10 dB has deviated to 0.1 GHz, 0.1 GHz, 0.1 GHz, and 0.05 GHz from the simulated results at their corresponding bands. These bandwidth characteristics satisfy the required application specifications. These measured bandwidth characteristics satisfy the Bluetooth, WiMAX, WLAN, fixed satellite/mobile specifications. Hence, the suggested antenna fulfills the requisite to multiband applications. The radiation patterns of quad band design shown in Fig. 6 exhibit nearly omnidirectional patterns in the H-plane and bidirectional patterns in the E-plane.

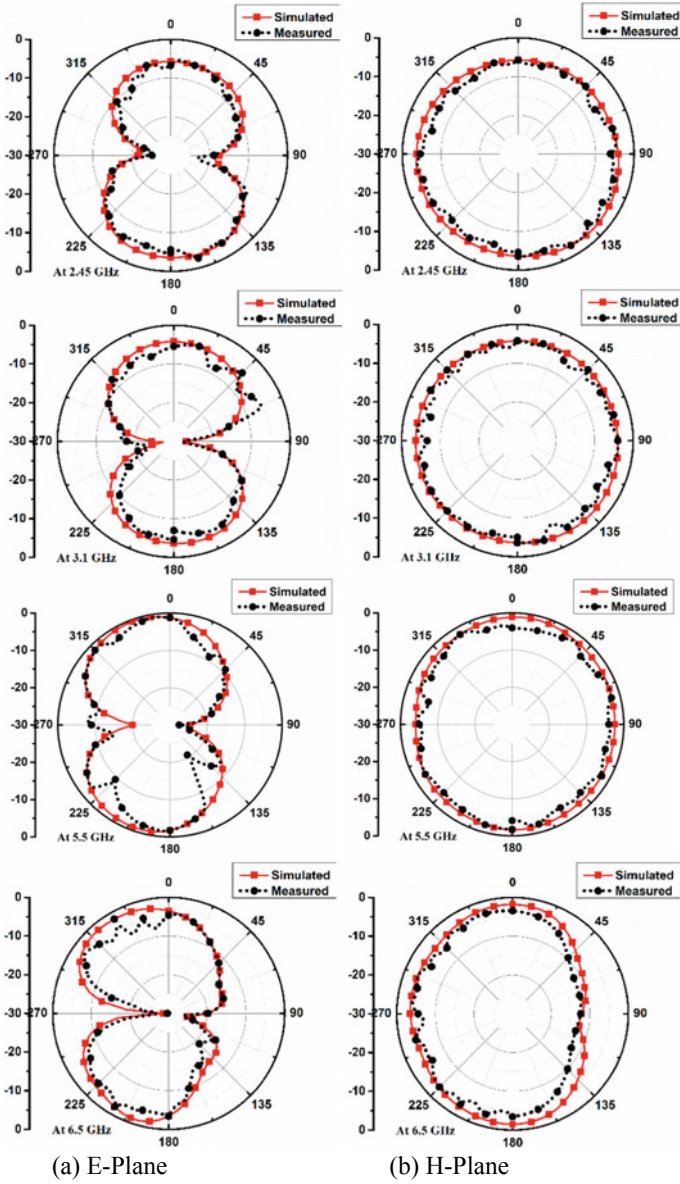


Fig. 6 Simulated and measured radiation pattern of Antenna #3

4 Conclusions

A simple quad band antenna has been designed and fabricated by creating open-ended slots on a rectangular radiating structure. By creating two rectangular open-ended and meandered slots on the radiator form a stub and resonates at four different frequencies. The proposed antenna can attain enough impedance bandwidth to cover Wi-Fi/Bluetooth, WiMax, WLAN, and fixed satellite operating bands with constant radiation pattern. Finally, the antenna is validated with measurement by the fabricated prototype. Also, it can be concluded the antenna is more compact, which has benefits of easy incorporation into printed circuit boards (PCBs).

References

1. J.F. Huang, M.T. Wu, J.Y. Wen, A compact triple-band antenna design for UMTS WLAN and WiMAX applications. *Microwave Opt. Technol. Lett.* **51**(9), 2207–2212 (2009)
2. H.W. Liu, H. Ku, Novel planar triple-band monopole antenna for WiMAX/WLAN applications. *Microwave Opt. Technol. Lett.* **52**(11), 2405–2408 (2010)
3. R. Kumar, Y. Kamatham, Compact tri-band monopole antenna for 2.4/5.8 GHz WLAN and 8.3 GHz ITU-T applications, in *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Coimbatore, India (2019), pp. 1–4
4. U. Chakraborty, S. Chatterjee, S.K. Chowdhury, P.P. Sarkar, A compact microstrip patch antenna for wireless communication. *Prog. Electromagnet. Res. C* **18**, 211–220 (2011)
5. K.G. Thomas, M. Sreenivasan, Compact triple band antenna for WLAN/WiMAX applications. *Electron. Lett.* **45**(16), 811–813 (2009)
6. A.Z. Manouare, S. Ibnyaich, A.E.L. Idrissi, A. Ghammaz, Miniaturized triple wideband CPW-fed patch antenna with a defected ground structure for WLAN/WiMAX applications. *J. Microwaves Optoelectron. Electromagn. Appl.* **15**(3) (2016)
7. J. Cao, X. Zhao, C. Liu, L. Yan, A planar compact triple-band monopole antenna for wlan/wimax applications. *Prog. Electromagn. Res. Lett.* **29**, 15–23 (2012)
8. W. Hu, Y.-Z. Yin, P. Fei, X. Yang, Compact triband square-slot antenna with symmetrical L-strips for WLAN/WiMAX applications. *IEEE Antennas Wirel. Propag. Lett.* **10**, 462–465 (2011)
9. X. Yang, F. Kong, X. Liu, C. Song, A CPW-fed triple-band antenna for WLAN and WiMAX applications. *Radio Eng.* **23**(4) (2014)
10. R. Kumar, Y. Kamatham, Compact tri-band monopole antenna for 2.4/5.8 GHz WLAN and 8.3 GHz ITU-T applications, in *IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Coimbatore, India (2019)
11. N. Kumawat, S. Gurjar, A compact triple and quadruple band antenna for bluetooth/WiMAX/WLAN/X-band applications, in *Proceedings of International Conference on ICT for Sustainable Development. Advances in Intelligent Systems and Computing* (Springer, Singapore, 2016), pp. 1–10
12. K. Yu, Y. Li, X.A. Luo, CPW-fed quad-band monopole antenna for L-band, WLAN and WiMAX communication applications, in *IEEE Progress in Electromagnetic Research Symposium (PIERS)*, Shanghai, China (2016), pp. 3996–4000

13. A. Khan, S.I. Naqvi, F. Arshad, Y. Amin, H. Tenhunen, A compact quad-band CPW-fed planar resonator for multiple wireless communication applications. *Appl. Comput. Electromagn. Soc. J.* **32**(11), 1001–1007 (2017)
14. D. Sarkar, K. Saurav, K.V. Srivastava, S. Awasthi, A compact penta-band CPW-fed monopole antenna using LC resonator and interdigital capacitor, in *IEEE Applied Electromagnetics Conference (2015)*, pp. 18–21

Design of a HNL Index Guiding Silica PCF for SCG: Applications in Dental Optical Coherence Tomography System



Mohammad Kamrozzaman Kiron, Md. Masum Hossen, Nguyen Hoang Hai, and Md. Anwar Hossain

Abstract In this manuscript, we present the planning and numerical investigation of SCG (Supercontinuum Generation) on a microstructure HNL-PCF (Highly nonlinear Photonic Crystal Fiber). SiO_2 is the background material for our proposed HNL-PCF where air holes are embedded within the cladding contour to supervise light confinement. The fiber has shown prospective application on dental optical coherence tomography (OCT) by possessing one near-zero dispersion point at operating wavelength 1300 nm. Authors come to the conclusion that wavelengths ranging from 1.3 to 1.5 μm have a better purpose of use in OCT systems. we have used the FEM for numerical calculation and nonlinear Schrodinger equation is used for SCG which is solved by split-step Fourier method. The study demonstrated 90 nm broad spectral (10 dB bandwidth), for an 80 m long fiber, is obtained by propagating the sech^2 2.5 ps pulse width at an FWHM (full width at half maxima) while the facility of the input optical pulse is 80 W. On its account, the utmost longitudinal resolutions within the depth direction for dental OCT are found: for enamel 4.10 μm and for dentin 4.39 μm .

Keywords Supercontinuum · Coherence · Optical coherence tomography (OCT) · Finite element method (FEM) · Longitudinal resolution

M. K. Kiron (✉) · Md. M. Hossen
Green University of Bangladesh, Dhaka, Bangladesh
e-mail: kamrozzaman@eee.green.edu.bd

N. H. Hai
Hanoi University of Science and Technology, Hanoi, Vietnam

Md. A. Hossain
Bangladesh University of Business and Technology, Dhaka, Bangladesh

1 Introduction

Supercontinuum (SC) means the continuity of the light spectrum. This phenomenon produces a continuous and expanded output spectrum by passing a non-monochromatic narrow band intense light pulse through an HNL medium. The physical processes, which account for spectral broadening [1] are FWM (Four-Wave Mixing), SPM (Self-Phase Modulation), XPM (Cross-Phase Modulation), Inelastic Scattering, Kerr nonlinearity, will be varied, counting on pump wavelength, fiber length peak power, pulse duration, and chromatic dispersion. Within the normal GVD (Group Velocity Dispersion) domain, SPM [2] and optical wave breaking [3] are chargeable for supercontinuum generation (SCG) [4]. After the primary pronouncement by Alfano and Shapiro in the 1970s, in a very bulky medium for the visible radiation source starting from 400 to 700 nm with the propagation of 5 mJ picosecond pulse, it grabbed researcher's attention to develop optimized light sources over conventional source [5, 6]. This spectral broadening occurrence opened a replacement possible medical application of these enhanced for identifying tiny biological entities like a personal cell or assessing subcellular structures nuclei by ensuring higher longitudinal resolution. Since HNL-PCFs contain a high ration contrast of refractive index between the core and cladding, it is been used frequently because of the medium for supercontinuum generation as compared to the standard optical fiber. Particularly, the look, flexibility of ratio contrast makes PCFs so popular for supercontinuum generation [7]. As silica HNL_PCF has some awe-inspiring properties like high birefringence [8], flatted chromatic dispersion [9], endlessly single-mode [10], high nonlinearity [11], large mode area [12], low confinement loss [13], its been using tremendously in optical coherence tomography [14], sensing [15], metagraphy, chromatography [16], spectroscopy [17], high power application [7], optical amplifier and switching [18], communication [19] for years. The higher the numerical aperture (NA) is essential for nondestructive OCT technique uses PCF-based SC light sources rather than superluminescent diodes (SLDs). For the OCT system, PCFs are designed in such a way that the SC can be generated in the normal dispersion regime which is important for the OCT imaging technique [1]. Previous studies showed OCT's applications with some potential advantages for biomedical and dental OCT system optimized with 1.050 μm wavelength [20] and 1.310 μm wavelength [1]. Though a high longitudinal resolution (2.5 μm at 1.300 μm) [1] was obtained, it became expensive as they used Keer-lense mode-locked Ti: Sapphire lasers for generating femtosecond pulse. SO as for all-embracing use in OCT systems researchers is attempting to develop some picosecond pulse laser pump source to scale back the value.

In this paper, we demonstrate an optimized HNL-Silica PCF, as SC generator, for OCT application at 1.3 μm wavelength where SCG occurs in normal group velocity dispersion (GVD) regime. We use a DFB (Distributed Feedback Laser) laser to get a picosecond pulse for the system [21]. Through this DFB, a sech^2 pulse containing 2.5 ps pulse width at FWHM (Full Width at Half Maximum) will be propagated through the proposed PCF. The output is observed by the split-step Fourier method

[22]. We have generated SC at $1.3 \mu\text{m}$ wavelength and calculated the coherence length of the generated SC is $6.77 \mu\text{m}$. We have also calculated the longitudinal resolutions in the depth direction in dental OCT which are $4.10 \mu\text{m}$ for enamel and $4.39 \mu\text{m}$ for dentin, because we considered the refractive index for enamel and dentin, respectively, 1.65 and 1.45.

2 Design Procedures of HNL-Silica PCF

A highly nonlinear index guiding PCF is proposed with all the optimize adjusted air holes, having diameters of d_1 , d_2 , d_3 , d_4 , and d_5 where silica (SiO_2) is employed as background material for the complete contour of the fiber. The fiber holds a solid high-index core and also the pitch of the PCF is Λ . The hexagonal structure is created effectively by the five-ring arrangement of embedded microstructure air holes within the cladding layer [23]. 5-degree structural parameters are used to optimized the design of the proposed HNL index guiding silica PCF. Pitch Λ , the diameter of the first ring of air holes and the diameter of the third ring of air holes, is kept very small, whereas the diameter of the second ring of air holes is kept comparatively large. The authors proved that small core PCFs are greatly leaky [21] (Fig. 1).

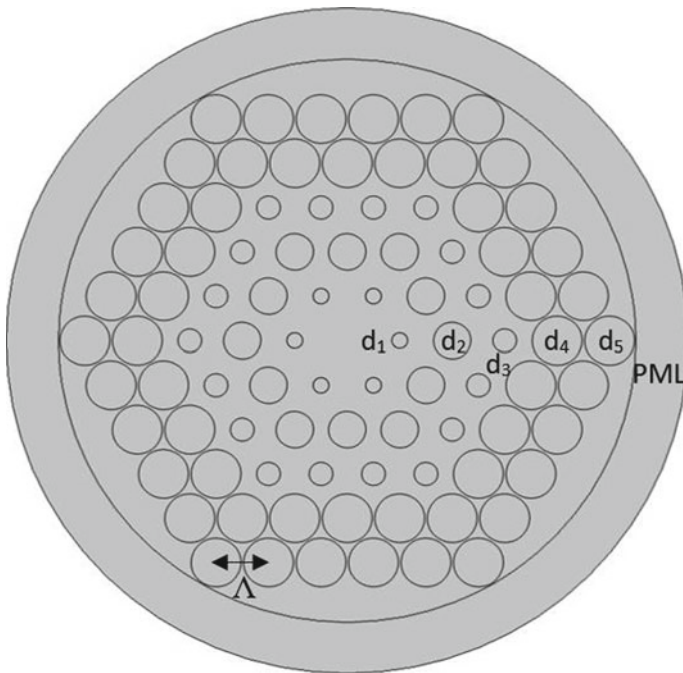


Fig. 1 Cross section of the silica PCF with optimal geometrical parameters

Confinement loss, chromatic dispersion and effective mode area play a vital role in designing HNL-PCF. Confinement loss needs to be reduced by keeping the air filling fractions, d/Λ large. We design the PCF in such a way so that the chromatic dispersion remains close to zero at the operating wavelength because the controlling chromatic dispersion is a very important issue for the real-life application of nonlinear optics. We were very careful while designing the PCF to keep the effective mode area very small because the effective mode area has an inverse relation with nonlinearity, which we want to have a large value because the larger values of nonlinearity will increase the SCG. The design process we follow is described in Refs. [24, 25], and the optimization of the design is described in Ref. [25]. At the beginning of the design, we kept all the air holes' diameter at their higher values then we gradually decreased to their optimum values. As first ring d_1/Λ plays the preeminent role on the dispersion curve, it is examined from several iterations by investigating the dispersion curve. Fine-tune of the dispersion curve was performed by varying the second ring of air holes d_2/Λ and the third ring of air holes d_3/Λ . d_4/Λ and d_5/Λ are kept at their maximum values to reduce the confinement loss.

3 Simulation Results and Discussions

To simulate the numerical properties of the HNL index guiding silica PCF FEM (Finite Element Method) is employed, which is one in every of the famous numerical tools having the handling capability of the outsized class of differential equations. Optimized structural parameters played the key role for the designed HNL-PCF to get the specified dispersion profile, for which spectrally broad SC is occurring at 1.30 μm pump wavelength. PML is imposed at the simulation contour, which is the most promising absorption edge precondition and helps the calculation to analyze the optical parameters of the proposed HNL-PCF [26, 27]. A hexagonal contour is created through a briefing of air holes of the triangular structured lattice. The radius of the inner ring of the PML is 4.785 μm where all 5 rings of air holes are placed and the radius of the outer circle of the PML is 5.655 μm ; therefore, the thickness of the PML is 0.87 μm . A ton of simulations has been performed to find the optical parameters of the proposed HNL-PCF through FEM. The process starts with the finding of the electric field $E(x, y)$ and effective mode index n_{eff} , after that dispersion parameter $D(\lambda)$, effective area A_{eff} , confinement loss L_c , nonlinear coefficient γ and GVD β_2 are calculated by using the subsequent equations [28, 29]:

$$D(\lambda) = -\frac{\partial^2 \text{Re}(n_{\text{eff}})}{\partial \lambda^2} \left(\frac{\lambda}{c} \right) \quad (1)$$

$$A_{\text{eff}} = \frac{(\iint_{-\infty}^{\infty} |E(x, y)|^2 dx dy)^2}{(\iint_{-\infty}^{\infty} |E(x, y)|^4 dx dy)} \quad (2)$$

$$L_c = 8.686 \times \frac{2\pi}{\lambda} \times \text{Im}(n_{\text{eff}}) \tag{3}$$

$$\gamma = \frac{2\pi n_2}{\lambda A_{\text{eff}}(\lambda)} \tag{4}$$

$$\beta_2 = \frac{\lambda^2}{2\pi c} D(\lambda) \tag{5}$$

where $\text{Re}(n_{\text{eff}})$ and $\text{Im}(n_{\text{eff}})$ are the real and imaginary part of the n_{eff} , $E(x, y)$ is the electric field, λ is the wavelength, c is the light velocity in vacuum. Air holes make the high-index contrast between core and cladding which helps the light to be confined inside the silica core. A small effective area makes the nonlinear coefficient higher by stiffening nonlinearity which helps the generated SC become broaden. Here in Eq. (4), n_2 is the nonlinear refractive index and the value of n_2 for pure silica is $2.1 \times 10^{-20} \text{ m}^2/\text{W}$ [1] and nonlinear coefficient $\gamma = 1 \text{ W}^{-1} \text{ km}^{-1}$ [24].

Parametric merit has a notable impact on the dispersion profile for PCFs. Figure 2 depicts the dispersion outline (dispersion, GVD, dispersion slope) of the recommended silica PCF with optimized parameters: for $\Lambda = 0.87 \mu\text{m}$, $d1 = 0.30 \Lambda$, $d2 = 0.71 \Lambda$, $d3 = 0.45 \Lambda$, $d4 = d5 = 0.95 \Lambda$. At $1.30 \mu\text{m}$ wavelength, we have got observed dispersion for the PCF is $-0.2598 \text{ ps}/(\text{nm km})$, $\text{GVD} = 0.2329 \text{ ps}^2/\text{km}$

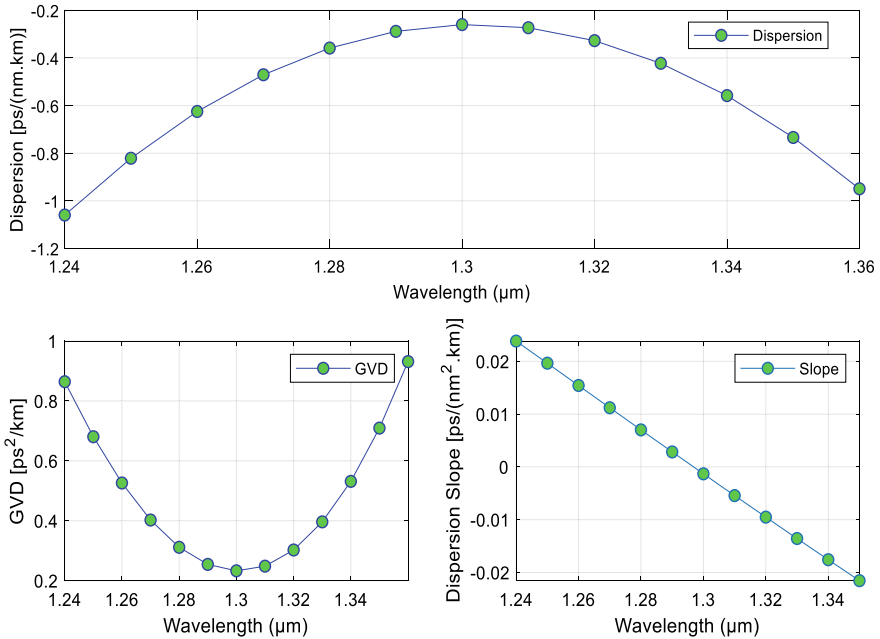


Fig. 2 Chromatic dispersion, group velocity dispersion and dispersion slope of the optimized HNL-PCF

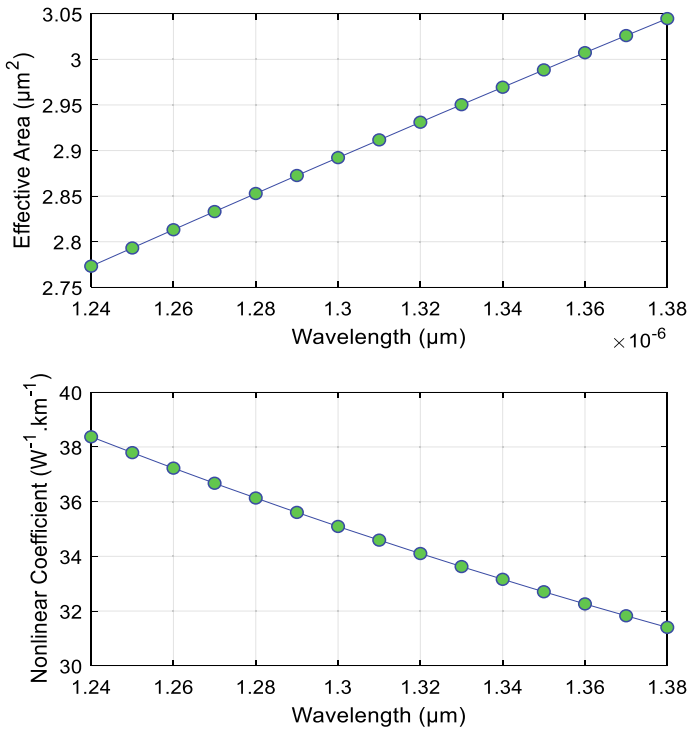


Fig. 3 Effective area and nonlinear coefficient of the profound HNL-PCF

and dispersion slope = $-0.001311 \text{ ps/nm}^2 \text{ km}$. Figure 3 illustrates influence on the effective area and nonlinear coefficient of the fiber, designed with optimized parameters, in relation to the wavelength (λ) in micrometer. The change has been observed within the wavelength range from in range from 1.24 to 1.38 μm . The graphs clearly illustrate that the effective area and nonlinear coefficient (γ) have shown an inverse response to wavelength (λ). While the effective area increases proportionally to wavelength, nonlinear coefficient (γ) falls-off with the increment in wavelength (λ). At 1.30 μm , we have found the effective area is about 2.892 μm^2 and the nonlinear coefficient (γ) is 35.09 $\text{W}^{-1} \text{ km}^{-1}$. Figure 4 shows the confinement loss (L_c) for the structured PCF on a graduated table.

Confinement loss (L_c) is calculated with reference to wavelength (λ) by using Eq. 3. By the accomplished arrangement of the air holes within cladding area, modes of confinement are often controlled and thus help us to urge a rational value of confinement loss (L_c) for designed PCF. We have found, at 1.30 μm wavelength, an economical value of confinement loss about only 1.979 dB/km.

Table 1 shows optimized parameters we've got accustomed design the HNL-Silica PCF. Pitch and therefore the first rings of air holes are decreased significantly to shape the dispersion curve, and therefore, the other parameters are accustomed to tuning

Fig. 4 Confinement loss of the optimized PCF

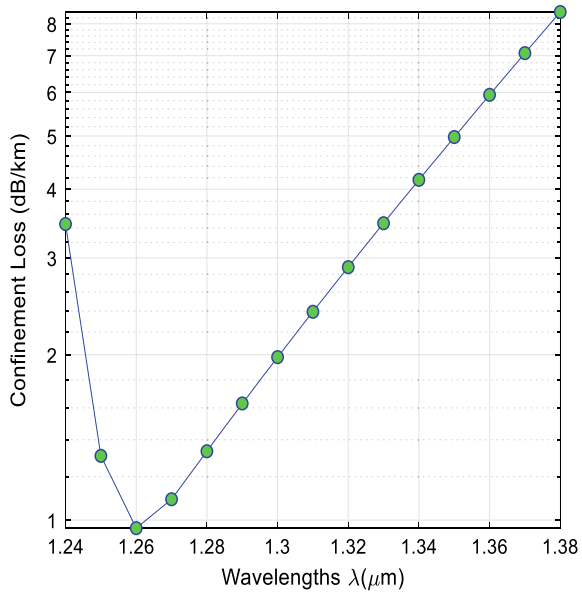


Table 1 Optimized structural parameter of the proposed PCF design

Parameters	Value
A	0.87 μm
d_1	0.30 A
d_2	0.71 A
d_3	0.45 A
d_4	0.95 A
d_5	0.95 A

the dispersion curve. The outer two rings of air holes are kept at their maximum value to cut back the confinement loss.

4 Supercontinuum Generation in HNL-Silica PCF

Nonlinear physical phenomena like Self-Phase Modulation (SPM), Four-Wave Mixing (FWM), Stimulated Raman scattering (SRS) is responsible for sustainable Supercontinuum Spectral Broadening. Here, we exhibit our findings which we observed from the numerical simulation for the proposed HNL-PCF within the picosecond regime. We have manipulated a distributed feedback laser diode as a pump source in this investigation. Nonlinear Schrodinger equation (NLSE) (Eq. 6) is solved by split-step Fourier method to offer rise to SC spectrum at 1.30 μm

wavelength, whereas the required higher-order dispersion parameters for NLSE are calculated through Taylor series (Eq. 7) [12, 29].

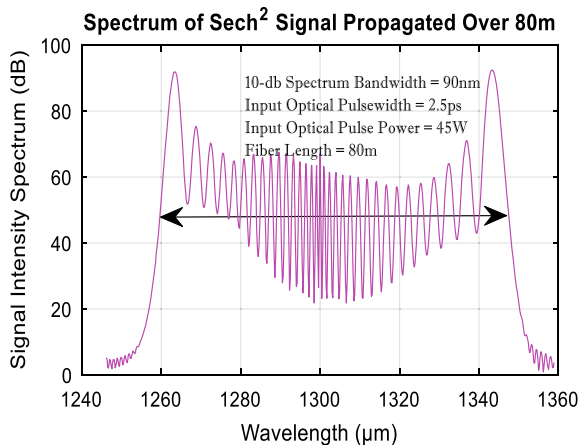
$$\frac{\partial A}{\partial z} + \frac{\alpha}{2}A + \frac{i}{2}\beta_2 \frac{\partial^2 A}{\partial T^2} - \frac{1}{6}\beta_3 \frac{\partial^3 A}{\partial T^3} - \frac{i}{24}\beta_3 \frac{\partial^3 A}{\partial T^3} = i\gamma \left[|A|^2 A + i \frac{\lambda_c}{2\pi c} \frac{\partial}{\partial T} (|A|^2 A) - T_R A \frac{\partial |A|^2}{\partial T} \right] \tag{6}$$

$$\beta(\omega) = n(\omega) \frac{\omega}{c} = \beta_0 + \beta_1(\omega - \omega_0) + \frac{1}{2}\beta_2(\omega - \omega_0)^2 \tag{7}$$

where A is a complex magnitude of the optical field, α is the attenuation coefficient of the fiber, γ is the nonlinear coefficient, T_R is Raman Scattering parameter, β is the propagation constant and $\omega_0 =$ center frequency and $\beta_i (i = 1-4)$ is the i th order of the Taylor series expansion of the propagation constant around the carrier frequency.

We have considered up to fourth-order dispersion effect to resolve NLSE by split-step Fourier method. Therefore, we adjusted the parameter’s value carefully for optimized fiber to generate SC because higher-order dispersion encompasses a consequential effect on heartbeat dynamics, particularly near the zero dispersion wavelength [30]. For the simulation, we have taken the parameters as nonlinear coefficient $\gamma = 35.09 \text{ W}^{-1} \text{ km}^{-1}$, $T_R = 3.0 \text{ fs}$, $\beta_2 = 0.2587 \text{ ps}^2/\text{km}$, $\beta_3 = -0.00143 \text{ ps}^3/\text{km}$, $\beta_4 = 0.0003265 \text{ ps}^4/\text{km}$ at the center wavelength at $\lambda_c = 1.30 \text{ }\mu\text{m}$. To urge the frequency contents of the spectrum, the resulted pulse is investigated through Fourier analysis. Figure 5 shows the power-dependent supercontinuum spectrum progressed through the designed PCF. At $1.30 \text{ }\mu\text{m}$ wavelength, a 90 nm 10-dB spectrum bandwidth is obtained employing a 80 m fiber length where a pulse of 45 W having 2.5 ps PW (pulse width) is provided from a pump source. OCT systems, the determination of intensity within the depth direction is crucial in order to maintain a specific degree of coherence for propagation distance coherent length (l_c). The

Fig. 5 Supercontinuum spectrum generation through HNL-silica PCF



coherent length (l_c) and axial resolution (l_r) in biotic substances are often estimated through the subsequent equations [31]:

$$l_c = \frac{2 \ln 2}{\pi} \frac{\lambda_c^2}{\nabla \lambda} \quad (8)$$

$$l_r = \frac{l_c}{n_{\text{tissue}}} \quad (9)$$

Here,

- n_{tissue} Refractive index of the living cell/tissue.
 $\nabla \lambda$ Bandwidth of the generated SC spectrum.
 λ_c Central wavelength.

Proposed HNL-Silica PCF has shown its feasibility toward dental OCT system: for enamel and dentin particularly, as they typically have a refractive index of n_{enamel} is 1.65 and n_{dentin} is 1.54, respectively [22]. We have found the coherent length (l_c) of 6.77 μm and maximum longitudinal resolution l_r of the SC light for enamel and dentin which are: for enamel $l_r = 4.10 \mu\text{m}$ and for dentin $l_r = 4.39 \mu\text{m}$.

5 Conclusion

A microstructured HNL-Silica PCF is meant to apply to the dental tomography system. It has forthcoming feasibility as we have got observed a little dispersion of $-0.2598 \text{ ps}/(\text{nm km})$, $\text{GVD} = 0.2329 \text{ ps}^2/\text{km}$ and dispersion slope = $-0.001311 \text{ ps}/\text{nm}^2 \text{ km}$ for the optimized parameters. For a solid silica core PCF with optimized parameters, we found nonlinear coefficient (γ) is about $35.09 \text{ W}^{-1} \text{ km}^{-1}$. A 90 nm 10-dB spectrum bandwidth employing a 80 m fiber length is found with a coherent length of 6.77 μm for a sech^2 input optical laser light pulse, having 2.5 ps pulse width at an FWHM and 45 W. Consequently, we found the utmost axial resolutions for dental tomography system which are 4.10 μm and 4.39 μm for enamel and dentin, respectively.

References

1. M.A. Hossain, Y. Namihira, S.M.A. Razzak, M.A. Islam, J. Liu, S.F. Kaijage, Y. Hirako, Design of all-normal dispersion highly nonlinear photonic crystal fibers for supercontinuum light sources: applications to optical coherence tomography systems. *Opt. Laser Technol.* **44**(4), 976–980 (2012)
2. L.E. Hooper, P.J. Mosley, A.C. Muir, W.J. Wadsworth, J.C. Knight, Coherent supercontinuum generation in photonic crystal fiber with all-normal group velocity dispersion. *Opt. Express* **19**(6), 4902–4907 (2011)

3. C. Finot, B. Kibler, L. Provost, S. Wabnitz, Beneficial impact of wave-breaking for coherent continuum formation in normally dispersive nonlinear fibers. *JOSA B* **25**(11), 1938–1948 (2008)
4. J.M. Dudley, G. Genty, S. Coen, Supercontinuum generation in photonic crystal fiber. *Rev. Mod. Phys.* **78**(4), 1135 (2006)
5. D. Huang, E.A. Swanson, C.P. Lin, J.S. Schuman, W.G. Stinson, W. Chang, M.R. Hee, T. Flotte, K. Gregory, C.A. Puliafito, J.G. Fujimoto, Optical coherence tomography. *Science* **254**, 1178–1181 (1991)
6. M.L. Gabriele, G. Wollstein, H. Ishikawa, L. Kagemann, Xu. Juan, L.S. Folio, J.S. Schuman, Optical coherence tomography: history, current status, and laboratory work. *Invest. Ophthalmol. Vis. Sci.* **52**(5), 2425–2436 (2011)
7. A. Ferrando, E. Silvestre, P. Andrés, J.J. Miret, M.V. Andrés, Designing the properties of dispersion-flattened photonic crystal fibers. *Opt. Express* **9**(13), 687–697 (2001)
8. J. Broeng, D. Mogilevstev, S.E. Barkou, A. Bjarklev, Photonic crystal fibers: a new class of optical waveguides. *Opt. Fiber Technol.* **5**(3), 305–330 (1999)
9. P.S.J. Russell, Photonic-crystal fibers. *J. Lightwave Technol.* **24**(12), 4729–4749 (2006)
10. I. Kaminow, T. Li (eds.), *Optical Fiber Telecommunications IV-B: Systems and Impairments*, vol. 2 (Elsevier, Amsterdam, 2002)
11. D.V. Skryabin, F. Luan, J.C. Knight, P.S.J. Russell, Soliton self-frequency shift cancellation in photonic crystal fibers. *Science* **301**(5640), 1705–1708 (2003)
12. K. Saitoh, M. Koshiba, Highly nonlinear dispersion-flattened photonic crystal fibers for supercontinuum generation in a telecommunication window. *Opt. Express* **12**(10), 2027–2032 (2004)
13. K.M. Gundu, M. Kolesik, J.V. Moloney, K.S. Lee, Ultra-flattened-dispersion selectively liquid-filled photonic crystal fibers. *Opt. Express* **14**(15), 6870–6878 (2006)
14. F. Poletti, V. Finazzi, T.M. Monro, N.G.R. Broderick, V. Tse, D.J. Richardson, Inverse design and fabrication tolerances of ultra-flattened dispersion holey fibers. *Opt. Express* **13**(10), 3728–3736 (2005)
15. K. Saitoh, M. Koshiba, Full-vectorial imaginary-distance beam propagation method based on a finite element scheme: application to photonic crystal fibers. *IEEE J. Quantum Electron.* **38**(7), 927–933 (2002)
16. K.P. Hansen, Dispersion flattened hybrid-core nonlinear photonic crystal fiber. *Opt. Express* **11**(13), 1503–1509 (2003)
17. K. Saitoh, M. Koshiba, T. Hasegawa, E. Sasaoka, Chromatic dispersion control in photonic crystal fibers: application to ultra-flattened dispersion. *Opt. Express* **11**(8), 843–852 (2003)
18. T.-L. Wu, C.-H. Chao, An efficient approach for calculating the dispersions of photonic-crystal fibers: design of the nearly zero ultra-flattened dispersion. *J. Lightwave Technol.* **23**(6), 2055 (2005)
19. T. Matsui, J. Zhou, K. Nakajima, I. Sankara, Dispersion-flattened photonic crystal fiber with a large effective area and low confinement loss. *J. Lightwave Technol.* **23**(12), 4178–4183 (2005)
20. P.A. Keane, H. Ruiz-Garcia, S.R. Sadda, Clinical applications of long-wavelength (1,000-nm) optical coherence tomography. *Ophthalmic Surg. Lasers Imaging Retina* **42**(4), S67–S74 (2011)
21. H.-F. Liu, Y. Ogawa, S. Oshiba, T. Nonaka, Picosecond pulse generation from a 1.3 μm distributed feedback laser diode using soliton-effect compression. *IEEE J. Quantum Electron.* **27**(6), 1655–1660 (1991)
22. M. Ohmi, Y. Ohnishi, K. Yoden, M. Haruna, In vitro simultaneous measurement of refractive index and thickness of biological tissue by the low coherence interferometry. *IEEE Trans. Biomed. Eng.* **47**(9), 1266–1270 (2000)
23. S.M.A. Razzak, Y. Namiyara, Tailoring dispersion and confinement losses of photonic crystal fibers using hybrid cladding. *J. Lightwave Technol.* **26**(13), 1909–1914 (2008)
24. V. Finazzi, T.M. Monro, D.J. Richardson, Small-core silica holey fibers: nonlinearity and confinement loss trade-offs. *JOSA B* **20**(7), 1427–1436 (2003)

25. T.-L. Wu, C.-H. Chao, A novel ultra-flattened dispersion photonic crystal fiber. *IEEE Photonics Technol. Lett.* **17**(1), 67–69 (2004)
26. S. Guo, F. Wu, Z. Albin, H. Tai, R.S. Rogowski, Loss and dispersion analysis of microstructured fibers by the finite-difference method. *Opt. Express* **12**(15), 3341–3352 (2004)
27. H. Rogier, D.D. Zutter, Berenger and leaky modes in optical fibers terminated with a perfectly matched layer. *J. Lightwave Technol.* **20**(7), 1141 (2002)
28. M. Chen, S. Xie, New nonlinear and dispersion flattened photonic crystal fiber with low confinement loss. *Opt. Commun.* **281**(8), 2073–2076 (2008)
29. G.P. Agrawal, Nonlinear fiber optics, in *Nonlinear Science at the Dawn of the 21st Century* (Springer, Berlin, 2000), pp. 195–211
30. B.G. Bale, S. Boscolo, K. Hammani, C. Finot, Effects of fourth-order fiber dispersion on ultrashort parabolic optical pulses in the normal dispersion regime. *JOSA B* **28**(9), 2059–2065 (2011)
31. I. Hartl, X.D. Li, C. Chudoba, R.K. Ghanta, T.H. Ko, J.G. Fujimoto, J.K. Ranka, R.S. Windeler, Ultrahigh-resolution optical coherence tomography using continuum generation in an air-silica microstructure optical fiber. *Opt. Lett.* **26**(9), 608–610 (2001)
32. F. Poli, A. Cucinotta, S. Selleri, A.H. Bouk, Tailoring of flattened dispersion in highly nonlinear photonic crystal fibers. *IEEE Photonics Technol. Lett.* **16**(4), 1065–1067 (2004)

A Machine Learning Approach for the Land-Type Classification



Arvind Kumar and Tejalal Choudhary

Abstract Land-type classification is an essential aspect of resource planning, and correctly classifying a land type can play a crucial role in designing and executing efficient utilization of land types for agriculture and other purposes. In recent years, machine learning (ML) methods have become popular, shown significant improvements in their performance and have been applied in various domains. In this work, we have proposed an ML-based approach for efficient and accurate classification of land types. We extensively experimented with different ML methods such as decision tree (DT), support vector machine (SVM), random forests (RF) and K-nearest neighbour (KNN). The empirical results suggest that ML-based approaches are superior for land-type classification. It is also found that out of the different ML methods applied on Statlog Landsat dataset, SVM outperforms other methods and achieves 92% accuracy better than other state-of-the-art methods.

Keywords Land classification · Statlog · Landsat · Soil classification · GIS

1 Introduction

In recent years, machine learning methods have shown outstanding result in classifying large datasets. Remote sensing field has attracted significant attention of the ML community researchers, and considerable progress has been made in different domains. A variety of techniques such as deep neural networks, genetic programming, decision tree (DT), support vector machine (SVM), K-nearest neighbour (KNN) and other analysis methods is applied in various real-world applications for extracting relevant information [1, 2]. However, a deep neural network method demands high computational requirements compared to the traditional machine

A. Kumar (✉) · T. Choudhary

Department of Computer Science and Engineering, Bennett University, Greater Noida 201310, India

e-mail: arvind.jki@gmail.com

T. Choudhary

e-mail: tejalal.choudhary@gmail.com

learning methods [3]. Remote sensing field has many applications such as land monitoring, estimating the damage in environments, radiation monitoring, urban planning, crop monitoring, soil quality evaluation and monitoring of the radiation to name a few. One of the essential parts of the remote sensing is the classification of land types. In the modern era of developed society, natural resource planning has emerged as a necessary and vital aspect of society. Land-type classification based on satellite imagery is one example of this. Data classification is the method of sorting and categorizing data into various distinct groups or distinct classes [4]. In other words, in machine learning, classification is a process, in which a sample instance is mapped to one of predefined labels or classes. The model learns patterns from the training data so that the acquired knowledge effectively maps the original samples to the proposed class or label without any help or assistance from a human professional. Thus, classification is a supervised machine learning technique that facilitates the categorization of an unseen sample into one of the subcategories present in the training dataset. As the model's training data knows the categories in advance in targeted class labels, this approach is called supervised learning.

Different machine learning (ML) techniques are actively applied in the natural resource planning field. These ML-based approaches not only save time but cost and resources too. Statlog dataset, also known as Landsat Satellite dataset, is publicly available at UCI repository [5]. A total of seven types of land exists in this dataset. The dataset contains different multispectral values in satellite imagery. The dataset has multispectral band pixel values in satellite imagery. The significant complexity of this dataset is its non-normal class distributions. Also, these classes are not well separated [6, 7]. In this work, we have proposed an efficient approach for the classification of land types. In particular, we have worked with different ML techniques to classify and evaluate other methods. A summary of the various popular machine learning (ML) algorithms, namely DT, SVM, KNN and random forests (RF), is given, and then classification has been done for this dataset. We evaluated different ML techniques to classify land type on Landsat dataset. The methods are evaluated using various evaluation metrics. The experimental results prove that the proposed approach can classify Landsat dataset with reasonable accuracy, eliminating human intervention.

This paper has been organized in various sections. In Sect. 2, we have discussed the existing contribution made by the researchers. Section 3 gives the details of the proposed methodology for the classification of UCI Statlog dataset and the details of the dataset and description of different ML techniques used in this work. Section 4 summarizes experimental points and evaluation metrics used. Finally, results and conclusion are given in Sects. 5 and 6, respectively.

2 Related Work

This section has discussed the research work done by other researchers in classifying the land-type datasets. Land-type classification is a popular field, and various

techniques have been proposed for its classification. The early works on classifying the soil type were by Vamanan and Ramar [8]. The authors proposed a data mining-based approach for the classification of soil types. Tej and Holban [9] uses an artificial neural network (ANN) and shows the highest accuracy of 82.56%. In their research, the authors find the optimal architecture of the neural network using clustering. In another study, Shine and Carr [10] classify the Statlog Landsat dataset. Their MATLAB implementation showed an accuracy of 89.8% for the KNN and 86% for the SVM. In one of the research [11], the authors proposed a new fuzzy k-means clustering algorithm to classify the Statlog Landsat dataset. The proposed method performed superior to the conventional fuzzy k-means clustering algorithm. Besides [12] also proposed using a spiking neural network with two different perturbations, namely sinusoidal and Gaussian. The authors applied a proposed approach for classifying Statlog Landsat dataset and two other datasets. Tayeb and Fizazi [13] proposed a two-step neural network-based classification approach. In the first step, the feature extraction is done on Statlog dataset using the proposed extractor multilayer perceptron (EMLP). In the second step, the results obtained in step 1 are classified using the SVM algorithm. The authors claim that their method outperforms other state-of-the-work methods. In another research [14], the authors analyse and compare SVM ensembles with four different ensembles techniques. The authors applied the proposed method on distinct datasets, including the Statlog Landsat. The MAdaBoostSVM gives an accuracy of 88.05 with ten base SVM classifiers and RBF kernel. In research [15], the authors proposed an incremental conic functions (ICF) algorithm based on mathematical programming to solve the classification problem. The proposed algorithm has improved results over previous version ICF-based classifier. The authors applied the proposed algorithm on the Statlog dataset. The propose algorithm is almost three times faster than the k-means-based polyhedral conic functions (PCF) algorithm without loss of classification accuracy.

3 Methodology

In this section, we have discussed our proposed methodology. Our methodology involved ensembles of ML approaches. This section first addressed the dataset used for the experiments, followed by a brief overview of the different ML methods and how we use them to classify land types.

3.1 Dataset Used

UCI land dataset is a popular dataset used to classify land types; it contains a dataset of seven land types. The dataset contains different multispectral band values in satellite imagery. Table 1 summarizes the details of various classes available in the dataset with their percentages. Figure 1 shows the distribution of the training and the test

Table 1 Statlog land-type dataset details

Id class	Training samples	Training %	Testing samples	Testing %
1 Red soil	1072	24.17	461	23.05
2 Cotton crop soil	479	10.80	224	11.20
3 Grey soil	961	21.67	397	19.85
4 Damp grey soil	515	9.36	211	10.55
5 Soil with vegetation stubble	470	10.60	237	11.85
6 Mixture class soil	0	0.00	0	0.00
7 Very damp grey soil	1038	23.40	470	23.50
Total	4435	100.0	2000	100.00

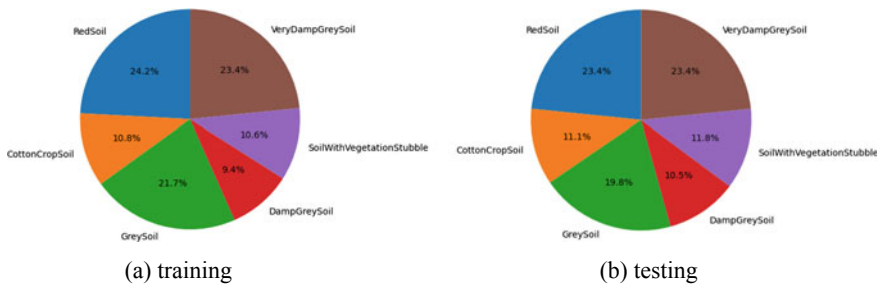


Fig. 1 Statlog (Landsat satellite) dataset

dataset, respectively. In Statlog (Landsat Satellite) dataset seven different types of categories are available: red soil (1), cotton crop soil (2), grey soil (3), damp grey soil (4), soil with vegetation stubble (5), mixture class soil (6) and very damp grey soil (7). The number in the parenthesis is a code for the respective class. The dataset has the multispectral values of pixels in 3×3 neighbourhoods in a satellite image and the categorization associated with each neighbourhood's central pixel.

The main aim is to classify land types for the given sample. The class of a pixel is represented as a number in the sample dataset. The dataset is a tiny subarea of a segment, which consisting of 82×100 pixels. Each line of data coincides with a 3×3 square neighbourhood of pixels entirely confined within the 82×100 subsegment. Each row holds the pixel values in the four spectral bands (which are transformed to ASCII) of each of the nine pixels in the 3×3 surroundings and a number representing the central pixel's classification tag. Thus, there are 36 columns as independent variables ($4 \times 3 \times 3$) and one output variable (class) in the dataset. All the attributes are numeric in the range [0–255]. Thus, the data can be categories into seven soil types. The dataset has been divided into two parts: training the model and others for the testing purpose. The training data have 4435 samples, with 24.17%, 10.8%, 21.67%, 9.36%, 10.6%, 0.0% and 23.4% of class 1, 2, 3, 4, 5, 6 and 7, respectively. Similarly, the testing data have 2000 points, with 23.05%, 11.20%,

19.85%, 10.55%, 11.85%, 0.00% and 23.50% percent of class 1, 2, 3, 4, 5, 6 and 7, respectively. It can be seen from the above numbers that data are unbalanced regarding different soil-type classes. The unbalanced dataset classification has its complexity, and in this case, classifiers give biased results towards the classes with higher sample counts [16].

3.2 Machine Learning Modelling Techniques

One of the vital application areas in machine learning domain is the classification, and it is a supervised learning technique. In classification, a model learns from the training data given as input to learning and then uses this acquired knowledge to classify or predict new samples or observations. Popular classification algorithms in machine learning are SVM, KNN, RF and DT.

Decision tree (DT): Decision trees are built using recursive partitioning to classify the data [17]. The algorithm chooses the most predictive feature to split the data one by one. In DT, the data are split depending upon the feature, and it is essential to decide “which attribute is the best, or more predictive”. Decision trees are created by dividing the training set into distinct nodes, in which one node consists of all, or the utmost, one category of the data [18]. Thus, DTs are about testing an attribute based on the test’s outcome, branching the cases. Each internal node coincides with a test, and each branch coincides with a result of the test. Each leaf node assigns a sample to a category or a class. A decision tree is constructed by considering the attributes one by one. This method may be summarized as the following:

1. The first step is to select an attribute from the training data.
2. The significance of the attribute in the splitting of the data is computed next. This is done to find whether the selected feature is useful or not (highest information gain).
3. Next, split the data based on the best attribute’s value.
4. Do this for each branch and repeat it for the remaining attributes.

K-nearest neighbour (KNN) Another supervised learning algorithm is K-nearest neighbour (KNN) [19]. In KNN, the samples are classified depending on their likeness to other instances in the dataset. The closer data points are known as “neighbour”. KNN is based on the assumption that the cases whose labels are same, they are close/near to each other. The distance from one sample to another sample is called the degree of dissimilarity. There are multiple ways of calculating similarity or distance for calculating the similarity/distance metric, such as Euclidean distance. The main steps of the KNN are as follows:

1. Select the initial value of K .
2. Calculate the distance from the new case (holdout from all of the cases in the dataset).

3. Find the K samples in the given sample data for training that are “closest or nearest” to the unknown sample data point measurements.
4. Predict the class or label of the unknown data point using the most popular class or label from the KNN.

Choosing the correct value of K is essential. A small value of K generates an extremely complex ML model, which might result in the model’s over-fitting. The prediction process is not generic and does not work well with out-of-sample data. The data that are not part of the dataset and used to train the model are called out-of-sample data. The KNN method may not be used to predict the unknown samples in a trusted way. It should be noted that over-fitting is not good, as we aim to generalize the model that works for every type of data, not only for the given training data. In contrast, if we take a large value of K , such as $K = 20$, the model becomes highly generalized, i.e. under-fitted. KNN can also work with a continuous target variable. If used for continuous target variable, the KNN’s average target value is used to predict new samples.

Random Forests (RF): Random forests (RF) were first introduced in 1995 by Tin Kam Ho of Bell Labs, and it is a supervised learning algorithm [20]. In RF, we build several decision trees and then group them to get a more precise and stable prediction. In RF, only random subsets of the features are chosen by the algorithm for splitting a node. Thus, it is an ensemble method (grounded on the divide-and-conquer approach) of DTs generated on a randomly divided dataset [21]. This collection of DT classifiers is also acknowledged as the forest. The individual DTs are created using an attribute selection indicator like information gain ratio, information gain and Gini index for each feature. Each tree depends on an independent random sample. In a classification process, each DT votes and the most voted class is chosen as the result. RF takes the mean of all the predictions, which removes the classification biases, so it does not suffer from the over-fitting problem. However, RF is slow in generating predictions because it has several DTs. In summary, there are four main steps:

1. Select samples randomly from the dataset.
2. For each sample, create a decision tree and calculate a prediction output value by each decision tree.
3. Do voting for each predicted result.
4. Select the final prediction based on maximum votes as the final prediction result.

Support Vector Machine (SVM) algorithm: SVM is a supervised ML technique that can categorize cases by finding or generating a separator [22]. SVM mechanism is based on mapping input data to a high-dimensional feature space. The sample data points can be separated into different categories, even when the given data are not linearly separable. After that, a separator is created for the data. The data should be transformed in such a way that a separator could be generated as a hyperplane. This plane can be utilized to classify new or unknown cases. Therefore, the SVM algorithm creates an optimal hyperplane that categorizes unseen examples. The main advantages of SVM are that they are accurate in high-dimensional spaces and use

a subset of training sample points in the decision functions. These decision functions are called support vectors. SVM is also memory efficient. The support vector machine's disadvantage is that if the features count is higher than the samples count, it leads to over-fitting. Another disadvantage of SVM is that it does not directly give probability estimates required in most classification tasks. And finally, SVMs are not very computationally efficient if the dataset is huge.

4 Experiments

For different machine learning techniques, ML models are developed using Python programming language on Windows 10 machine. Python libraries NumPy, Pandas and SKLearn are used to implement the ML models. The dataset has two different sets: training (4435 records) and another for testing (2000 records). Various models are trained, tested and evaluation metrics are calculated. Accuracy, precision, recall and $F1$ -score metrics are computed for all seven classes and mentioned techniques.

4.1 Performance Evaluation

The popular evaluation metrics for classification problems are accuracy, precision, recall and $F1$ -score [23]. All the methods are evaluated on above-listed evaluation metrics for their effectiveness. The precision, recall, accuracy and $F1$ -score are given as:

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}} \quad (1)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

$$\text{Sensitivity/Recall} = \frac{\text{TP}}{\text{FN} + \text{TP}} \quad (3)$$

$$F1 \text{ score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

5 Results

This section provides the details of the various experiments performed with the selected dataset and different ML techniques. The results for all four considered

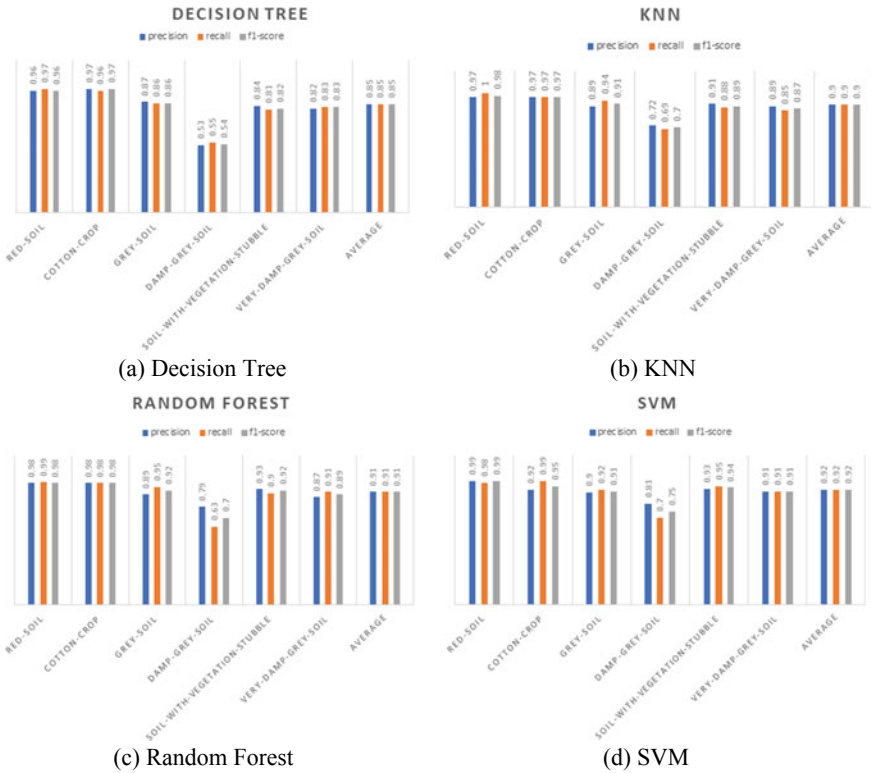


Fig. 2 Result summary

methods are shown in Fig. 2. In Fig. 2a–d shows the precision, recall and F1-score for the decision tree, K-nearest neighbour, random forests and support vector machine, respectively. It can be seen from Fig. 2a, c and d that red soil and cotton crop have the highest scores for the decision tree, random forests and SVM compared to other classes in the dataset. Figure 2b shows that red soil has the highest classification scores in K-nearest neighbour’s case.

A comparison of different accuracy achieved with each method is shown in Table 2. It can be seen from Table 2 that SVM outperformed all other ML techniques and

Table 2 Result summary

ML technique	Accuracy	Precision	Recall	F1-score
Decision tree	85.05	0.85	0.85	0.85
KNN	90.20	0.90	0.90	0.90
Random forest	91.20	0.91	0.91	0.91
SVM	92.00	0.92	0.92	0.92

Table 3 Comparison with other methods

Referenced method	Accuracy (%)
Tej and Holban [9]	82.56
Shine and Carr [10]	89.80
Wang et al. [14]	88.05
Tis Work (SVM)	92.00

has the best accuracy of 92%. It is important to note that the decision tree performed worst with an accuracy of 85%. The best accuracies of KNN and random forests are 90 and 91, respectively.

Table 2 also summarizes the precision, recall and $F1$ -score for all the methods. It can be seen from Table 2 that the SVM has the highest accuracy precision, recall and $F1$ -score. Thus, for Statlog land-type data classification, SVM performed best with 92% average accuracy and decision tree techniques perform lowest with 85.05% of average accuracy. Table 3 shows the comparison of the proposed ML-based approached to classify land type with other approaches.

6 Conclusion

Land classification or soil classification is one of the crucial problems of natural resource planning. In this work, Statlog Landsat dataset, which contains a dataset of seven land types, is taken from the UCI repository. Four popular machine learning algorithms, namely SVM, random forests, KNN and decision tree, classify land types. The performance of the various methods is evaluated by accuracy, recall, precision and $F1$ -score. The experimental results suggest that classifying land type with ML methods gives higher accuracies. We also found that SVM outperformed all other three techniques with an accuracy of 92%. In summary, ML-based land-type classification facilitates easy, faster and low-cost natural resource planning that can help in agriculture as well as in other areas.

References

1. G. Singal, A. Goswami, S. Gupta, C.T. Pitfree, Pot-holes detection on Indian roads using mobile sensors, in *2018 IEEE 8th International Advance Computing Conference (IACC)* (IEEE, 2018), pp. 185–190
2. A. Kumar, A. Kakkar, R. Majumdar, A.S. Baghel, Spatial data mining: recent trends and techniques, in *2015 International Conference on Computer and Computational Sciences (ICCCS)* (IEEE, 2015), pp. 39–43
3. T. Choudhary, V. Mishra, A. Goswami, J. Sarangapani, A comprehensive survey on model compression and acceleration. *Artif. Intell. Rev.* 1–43 (2020)
4. A. Kumar, N. Sinha, A. Bhardwaj, A novel fitness function in genetic programming for medical data classification. *J. Biomed. Inf.* **112**, 103623 (2020)

5. A. Srinivasan, UCI machine learning repository, statlog (landsat satellite) data set (2017). Available from: [https://archive.ics.uci.edu/ml/datasets/Statlog+\(Landsat+Satellite\)](https://archive.ics.uci.edu/ml/datasets/Statlog+(Landsat+Satellite))
6. A. Anand, L. Wilkinson, D.N. Tuan, An L-infinity norm visual classifier, in *2009 Ninth IEEE International Conference on Data Mining* (IEEE, 2009), pp. 687–692
7. A. Kumar, N. Sinha, Classification of forest cover type using random forests algorithm, in *Advances in Data and Information Sciences* (Springer, Singapore, 2020), pp. 395–402
8. R. Vamanan, K. Ramar, Classification of agricultural land soils a data mining approach. *Int. J. Comput. Sci. Eng.* ISSN. 2011:0975-3397
9. M.L. Tej, S. Holban, Determining neural network architecture using data mining techniques, in *2018 International Conference on Development and Application Systems (DAS)* (IEEE, 2018), pp. 156–163
10. J.A. Shine, D.B. Carr, A comparison of classification methods for large imagery data sets. *JSM* 3205–3207 (2002)
11. C.T. Chang, J.Z. Lai, M.D. Jeng, A fuzzy k-means clustering algorithm using cluster center displacement. *J. Inf. Sci. Eng.* **27**(3), 995–1009 (2011)
12. J. Yang, P. Zhang, Y. Liu, Robustness of classification ability of spiking neural networks. *Nonlinear Dyn.* **82**(1–2), 723–730 (2015)
13. M.S. Tayeb, H. Fizazi, A new neural architecture for feature extraction of remote sensing data. *Int. J. Comput. Sci. Eng.* **21**(1), 95–104 (2020)
14. S.J. Wang, A. Mathew, Y. Chen, L.F. Xi, L. Ma, J. Lee, Empirical analysis of support vector machine ensemble classifiers. *Expert Syst. Appl.* **36**(3), 6466–6476 (2009)
15. E. Cimen, G. Ozturk, O.N. Gerek, ICF: an algorithm for large scale classification with conic functions. *SoftwareX* **8**, 59–63 (2018)
16. L. Li, H. He, J. Li, Entropy-based sampling approaches for multi-class imbalanced problems. *IEEE Trans. Knowl. Data Eng.* **32**(11), 2159–2170 (2019)
17. S.R. Safavian, D. Landgrebe, A survey of decision tree classifier methodology. *IEEE Trans. Syst. Man Cybern.* **21**(3), 660–674 (1991)
18. J. Chen, J.K. Patel, R. Vasques, Solver recommendation for transport problems in slabs using machine learning. arXiv preprint [arXiv:190608259](https://arxiv.org/abs/190608259) (2019)
19. G. Shakhnarovich, T. Darrell, P. Indyk, *Nearest-Neighbor Methods in Learning and Vision: Theory and Practice (Neural Information Processing)* (The MIT Press, Cambridge, 2006)
20. T.K. Ho, Random decision forests. *Doc. Anal. Recogn.* **1**, 278–282 (1995)
21. P.O. Gislason, J.A. Benediktsson, J.R. Sveinsson, Random forests for land cover classification. *Pattern Recogn. Lett.* **27**(4) (2006)
22. X. Gu, T. Ni, H. Wang, New fuzzy support vector machine for the class imbalance problem in medical datasets classification. *Sci. World J.* **2014** (2014)
23. L. Cuadros-Rodríguez, E. Pérez-Castaño, C. Ruiz-Samblás, Quality performance metrics in multivariate classification methods for qualitative analysis. *TrAC Trends Anal. Chem.* **80**, 612–624 (2016)

A Tiny CNN Architecture for Medical Face Mask Detection for Resource-Constrained Endpoints



Puranjay Mohan , Aditya Jyoti Paul , and Abhay Chirania 

Abstract The world is going through one of the most dangerous pandemics of all time with the rapid spread of the novel coronavirus (COVID-19). According to the World Health Organization, the most effective way to thwart the transmission of coronavirus is to wear medical face masks. Monitoring the use of face masks in public places has been a challenge because manual monitoring could be unsafe. This paper proposes an architecture for detecting medical face masks for deployment on resource-constrained endpoints having extremely low memory footprints. A small development board with an ARM Cortex-M7 microcontroller clocked at 480 MHz and having just 496 KB of framebuffer RAM has been used for the deployment of the model. Using the TensorFlow-Lite framework, the model is quantized to further reduce its size. The proposed model is 138 KB post-quantization and runs at the inference speed of 30 FPS.

Keywords ARM Cortex-M7 · COVID-19 · Edge computing · Face mask detection · Quantization · SqueezeNet · TinyML

P. Mohan (✉) · A. J. Paul
Cognitive Applications Research Lab, Chennai, India
e-mail: puranjaymohan_mu@srmuniv.edu.in

A. J. Paul
e-mail: aditya_jyoti@srmuniv.edu.in

P. Mohan · A. Chirania
Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu 603203, India
e-mail: abhaychirania_de@srmuniv.edu.in

A. J. Paul
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu 603203, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_52

1 Introduction

The sudden increase of computational capability and availability of data in the last few years has allowed intelligent systems to solve various problems involving computer vision, speech, etc. Traditionally, these models got deployed on servers with high compute and storage capabilities. With the rise of the Internet of things and edge computing, the need to deploy these systems at the edge has grown. The major roadblock in edge deployment of deep neural networks is the very high computational and memory footprint of these models. Image classification is one such problem where edge deployment is in high demand because of the many applications which rely on it. Face mask detection is a subset of image classification where the goal is to classify the image into two classes, i.e., mask and no mask, respectively.

The outbreak of the novel coronavirus has significantly impacted the livelihood of people across the globe [1], and effective deployment of face mask detection in public places can help in thwarting the transmission of the virus. Face mask detection is a non-trivial problem because of its high throughput, reliability, and privacy requirements; it cannot use traditional deployment methods where the image is first sent to a server for classification and the result is sent back for further use in the application.

One application scenario the authors of this research work envision is a small camera attached to an automatic door, and the camera continuously takes images of the person standing in front of it and only opens if the person is wearing a mask. This intelligent door can be used at all public places, and it will safely monitor the people entering public premises. One requirement of this smart door is that it should be cheap and consume less energy. The proposed model is small enough to fit inside the memories of the smallest and cheapest microcontrollers available in the market. Many other applications requiring high-speed mask detection on the edge can be envisaged as other possible applications as well.

TinyML is an upcoming field at the intersection of hardware, software, and machine learning algorithms that is gaining massive traction. Recent developments in this field include building deep neural networks having sizes of few hundreds of KBs. This paper presents the process to train and deploy an innovative architecture on the OpenMV H7 development board for detecting face masks using the small on-board camera.

A major challenge with TinyML is that most microcontrollers do not have a floating-point unit, and hence, all mathematical computations need to work on integers. This leads to a smaller model along with a change in accuracy that is difficult to predict. Earlier studies focused on deployment on more powerful edge devices, using much larger models, but this paper reports how the quantized CNN model compares to the existing architectures in terms of model size, accuracy, and performance.

The rest of the paper is organized as follows—Sect. 2 covers the literature review, Sect. 3 discusses some technical and hardware details, Sect. 4 explains the experimental methodology, Sect. 5 reports the observations and the findings, and Sect. 6 concludes this paper, throwing some light on possible future research avenues in this field.

2 Literature Review

In this section, some prior advances in face mask detection and quantization, which are the primary facets of this research work, have been reviewed.

2.1 Face Mask Detection

Due to the coronavirus pandemic, face masks have become an integral part of our society; hence, numerous implementations of detecting face mask have come forward which are based on convolutional neural networks [2]. [3] proposed a two-stage detection scheme, the first being face detection, and the next being face mask classifier. Loey et al. [4] proposed a hybrid deep transfer learning model with two components, the first for feature extraction using ResNet50 and other for classification using SVM, and other ensemble algorithms. Their support vector machine (SVM) classifier achieved testing accuracy of 99.64%. RetinaMask [5] achieved state-of-the-art result on public face mask dataset (2.3 and 1.5% higher precision than baseline result on face and mask detection) using one-stage detector which consisted of feature pyramid network to fuse high-level semantic information with multiple feature maps. Meenpal et al. [6] presented a model with pre-trained weights of VGG-16 architecture for feature extraction and then a fully connected neural network (FCNN) to segment out faces present in an image and detect face masks on them. The model showed great result in recognizing non-frontal faces. Loey et al. [7] proposed a model using YOLO-v2 with ResNet-50 which achieves higher average precision by using mean Intersection over Union (IoU). Jignesh Chowdary et al. [8] proposed a model based on transfer learning with InceptionV3. It outperformed recently proposed models by achieving testing accuracy of 100% on simulated masked face dataset (SMFD). Roy et al. [9] discussed the challenge of implementing object detection on edge devices, and the paper compared various popular object detection algorithms like YOLO-v3, YOLO-v3tiny, Faster R-CNN, etc., to determine the most efficient algorithm for real-time detection of face masks.

2.2 Quantization

Leveraging quantization techniques are necessary for implementing CNNs on resource-constrained devices. Banner et al. [10] introduced 4-bit training quantization on both activation and weights, achieving accuracies, a few percent less than state-of-the-art baselines across CNNs. Nahshan et al. [11] proposed a method which quantizes layer parameters that improve accuracy over the existing post-training quantization techniques. Zhao et al. [12] proposed an outlier channel splitting (OCS)-based method to improve quantization performance without retraining. Choukroun et al. [13] discussed low-bit quantization of neural networks by optimization of con-

strained mean squared error (MSE) problems for performing hardware-aware quantization. Jacob et al. [14] proposed a quantization scheme along with a co-designed training procedure. The paper concluded that inference using integer-only arithmetic performs better than floating-point arithmetic on typical ARM CPUs. Gong et al. [15] proposed Differentiable Soft Quantization (DSQ) to bridge the gap between the full-precision and low-bit networks. The hybrid compression model in [16] uses four major modules, approximation, quantization, pruning, and coding, which provides 20–30x compression rate with negligible loss in accuracy. The research by Dong et al. [17, 18] and Yao et al. [19] proposed mixed-precision quantization techniques, where more sensitive layers are at higher precision.

3 Technical Details of the Problem

This section explains the technical details related to the experimental setup including the hardware, the software, and the use of datasets.

3.1 *Hardware Setup for Deployment*

The hardware used in this research for edge deployment is the OpenMV Cam H7 [20], housing STMicroelectronics' STM32H743VI [21], an ARM Cortex-M7-based 32-bit microcontroller, and a small camera. The microcontroller has a clock speed of 480 MHz, 1 MB SRAM for various applications and 2 MB of flash memory for non-volatile storage. The development board provides a MicroPython-based operating system allowing easy deployment and on-device analytics of TF-Lite models.

The documentation of the board suggests keeping the model under 400 KB, but during this study [22], we found that the biggest model which can be fit successfully in memory is under 230 KB. A larger model of size up to 1 MB can be stored on the flash memory, but for that, the model has to be converted into a FlatBuffer using STM32Cube.AI and the operating system have to be recompiled, which leads to the loss of the utility of the MicroPython.

The models were trained using Kaggle kernels with Tensor Processing Unit (TPU) acceleration enabled, 128 TPU elements per core with eight such cores running in parallel, and this made the training time extremely short.

3.2 *Dataset Construction*

Four datasets were used in this research work. The details of the datasets can be seen in Table 1. The first dataset from Kaggle [23] has around 11,792 images taken on different backgrounds and cropped to the face region. The images of this dataset were

Table 1 Dataset details

S. No	Name	Original	Augmented
1	Face mask 12K images dataset	11,792	58,960
2	Face mask classification	440	22,200
3	OpenMV dataset	1979	49,895
Total data for training and validation		14,211	131,055
Total data for testing		594	4794

**Fig. 1** Images from OpenMV Cam H7 Dataset

merged and interpolation augmentation was applied using OpenCV's interpolation methods, INTER_AREA, INTER_CUBIC, INTER_NEAREST, INTER_LINEAR, and INTER_LANCZOS4 to augment the images to 58,960.

The second dataset was also from Kaggle [24] which had 440 images taken on noisy backgrounds, equally divided into mask and without mask images. It was augmented to 22,200 using standard augmentation followed by interpolation.

The third dataset was produced by the authors using the OpenMV Cam H7 camera. Images of size 200×200 were taken and saved on the SD card of the development board. This dataset had 1979 images which were augmented to 49,895 using augmentation techniques discussed earlier. Some images from this dataset can be seen in Fig. 1.

A fourth dataset was also produced using the OpenMV camera which had 594 images augmented to 4794. This dataset was held out for testing the performance of the OpenMV Cam H7. The exact usage of this dataset is novel to this research work and has been elaborated in Sect. 4.

All the images of each dataset were resized to 32×32 as it was found to be the optimal size of the image to fit in the framebuffer of the microcontroller.

4 Experimental Methodology

This section explains the steps taken in this research work for data splitting, model design, evaluation, and model comparison.

4.1 Data Splitting

The datasets after being merged had 131,055 images in total, and these included images from two Kaggle datasets [23, 24] and one dataset produced by the OpenMV camera. The fourth dataset was held out and was used for testing.

This regime of holding out a separate dataset for testing is not usually performed but was considered imperative in this research work to evaluate the generalization achieved by the models running on the microcontroller. The normal regime of combining everything and then taking out a small portion for testing would be unable to show the true model performance on the target edge-case scenarios due to differences between the distributions of the train and test sets.

4.2 Proposed Architecture and Comparison with SqueezeNet

After experimenting with different architectures and comparing their size and performance, the CNN architecture shown in Fig. 2 was found to be the best, after considering the RAM constraints of the device. This model has 128,193 trainable parameters, and full integer quantization reduces it to 138 KB.

SqueezeNet [25] was chosen for comparison with the proposed model because of its small size. A smaller version of SqueezeNet was also built by removing two fire modules, which is called modified SqueezeNet in this work, and also used for comparison.

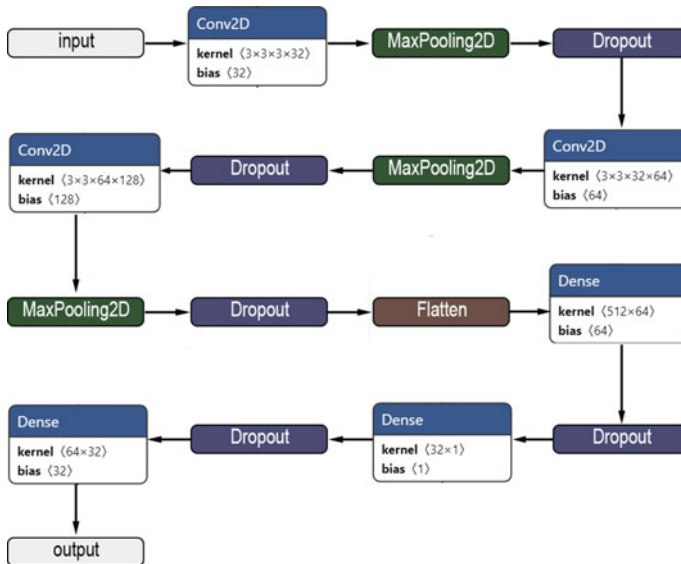


Fig. 2 Proposed architecture

4.3 Training Specifications

All three models were designed in TensorFlow 2.3. BinaryCrossentropy loss was chosen as the loss function, shown in Eq. (1), because the problem involved binary classification of images.

$$\text{Loss} = -(y \log(p) + (1 - y) \log(1 - p)) \quad (1)$$

where y is the true label and p is the predicted label.

Adam was chosen as the optimizer with a learning rate of 0.001, first moment decay of 0.9, second moment decay of 0.999, and epsilon was chosen as 10^{-7} . ReduceLROnPlateau callback was used to reduce the learning rate by a factor of 0.2 when the validation accuracy did not improve for five epochs. ModelCheckpoint callback was used to save the best weights into a file.

4.4 Post-training Procedure

TensorFlow-Lite's Full Integer Quantization was used to convert all three models from float32 precision to int8 precision. This procedure used a representative dataset for this conversion using the dynamic range of activations. This representative dataset was built by taking a small part of the test set.

4.5 Evaluation of the Quantized Models

All three models were evaluated using the TensorFlow-Lite Interpreter. The quantized models were loaded into the Interpreter, and OpenMV test set was used to find the classification metrics.

On-device evaluation was not possible for the SqueezeNet and the modified Squeezenet because both of them were bigger than 230 KB. The proposed model was loaded onto the OpenMV Cam; a script took images of size 200×200 , scaled them to 32×32 , and normalized them before feeding it to the model. All images and predictions were saved on the SD card and later used for the analysis.

5 Results and Discussion

This section explains the results and analysis, along with comparison of the proposed model against SqueezeNet and modified SqueezeNet.

5.1 SqueezeNet Model

The methodology discussed in sec. 4 was followed to train the SqueezeNet model. The training accuracy reached 99.79%, achieving a test accuracy of 98.50% for the float32 model and 98.53% for the int8 model. The size of the float32 model was 8 MB which shrunk to 780 KB post-quantization. The details can be seen in Fig. 3.

5.2 Modified SqueezeNet Model

Modified SqueezeNet was also trained in a similar way. Size of the float32 model was 3.84 MB which got reduced to 386 KB after quantization. The test accuracies for this model were 98.93 and 98.99% for float32 and int8, respectively. The details can be seen in Fig. 4.

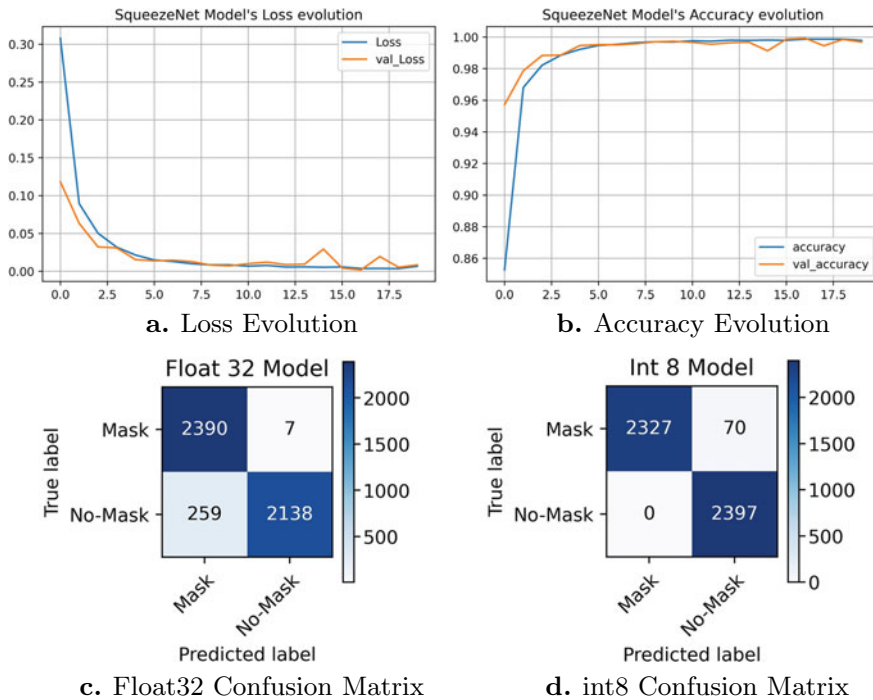


Fig. 3 SqueezeNet

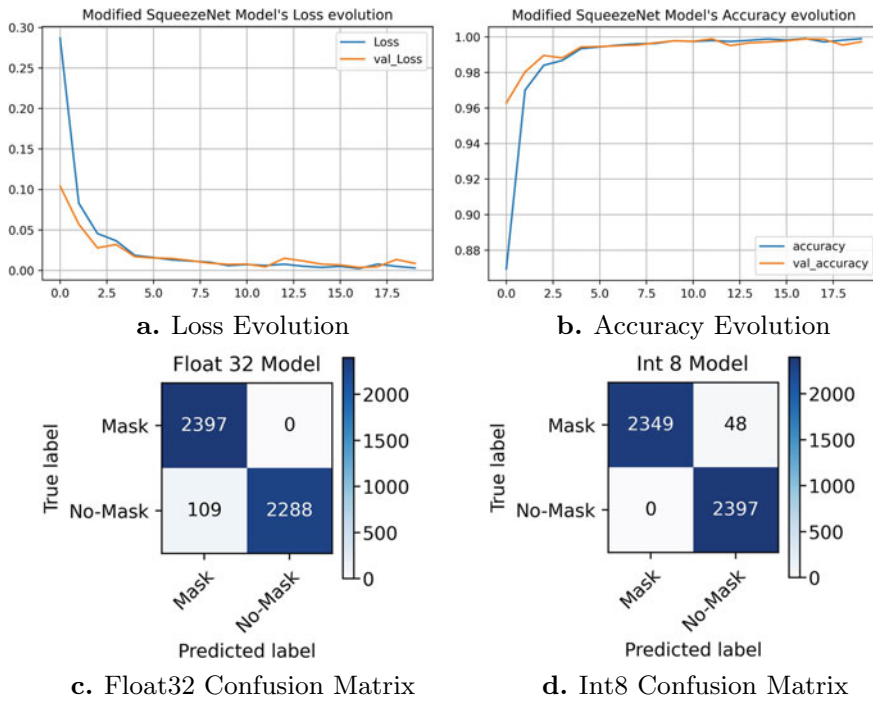


Fig. 4 Modified squeezeNet

5.3 Our Proposed Model

The proposed model reached the training accuracy of 99.79% and achieved a testing accuracy of 99.81% and 99.83% for float32 and int8 models, respectively. The 1.52 MB float32 model was reduced to 138 KB post-quantization. Our model outperformed SqueezeNet and modified SqueezeNet in both accuracy and size. The details can be seen in Fig. 5.

5.4 Discussion

On comparing the SqueezeNet and the modified SqueezeNet, it was observed that the modified version, which had two fire modules removed, generalized better than the original model, keeping the precision constant. Thus, it was observed,

“On resource-constrained endpoints, smaller models sometimes outperform bigger ones in generalizing to new data.”

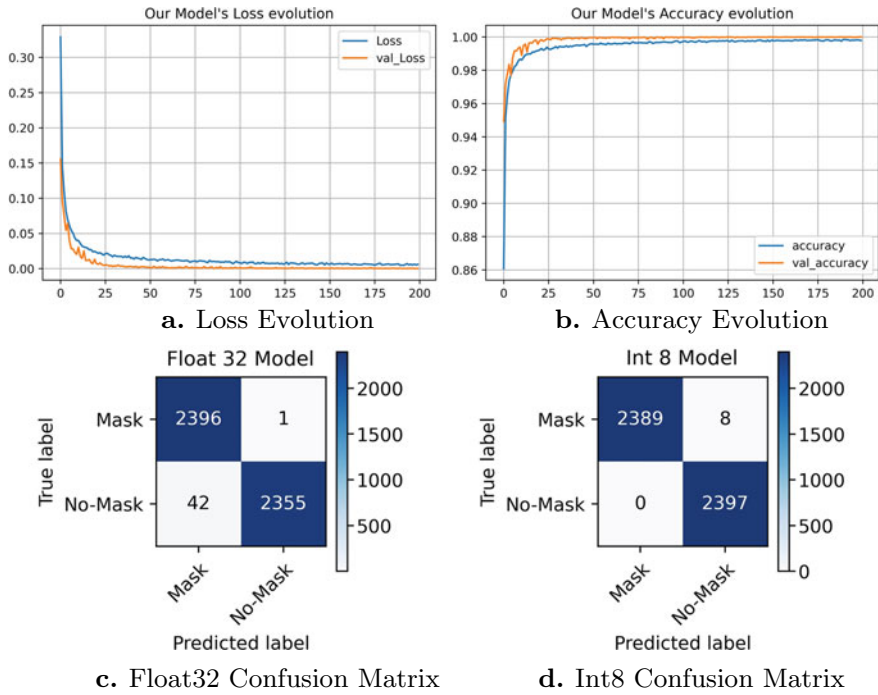


Fig. 5 Our model

It was also observed,

“On devices with Floating Point Unit (FPU) support, keeping inputs and outputs as float32 gives best results.”

The proposed model, despite being the smallest one, achieved the highest accuracy among all three. Since the int8 accuracy is a little more than float32 accuracy for all experimental models, the following conclusion can be drawn,

“Int8 appears to generalize better than float32 for small models.”

Table 2 shows the size and accuracy comparison of the three models. Table 3 illustrates the compression comparison. The proposed model, despite having the smallest size, achieved the highest accuracy, precision, recall, and F1 score, as can be seen in Tables 4 and 5, representing the int8 and float32 models, respectively. Some of the model predictions can be seen in Fig. 6. Even smaller CNNs may overfit when solving problems like binary classification; hence, aggressive regularization is required to increase their generalization accuracy.

Table 2 Model comparison

Model name	float32 size (MB)	float32 test acc. (%)	int8 size (KB)	int8 test acc. (%)
SqueezeNet	8.0	98.50	780.0	98.53
Modified squeezeNet	3.8	98.93	386.0	98.99
Our model	1.5	99.81	138.0	99.83

Table 3 Compression comparison

Model name	float32 → int8 reduction in size (KB)	Size reduction percentage (%)
SqueezeNet	7417	90.48
Modified squeezeNet	3546	90.18
Our model	1428	91.16

Table 4 int8 classification report

Label	Precision	Recall	F-1 score
<i>SqueezeNet</i>			
Mask	1	0.97	0.99
No-mask	0.97	1	0.99
<i>Modified squeezeNet</i>			
Mask	1	0.97	0.99
No-mask	0.97	1	0.99
<i>Our model</i>			
Mask	1	1	1
No-mask	1	1	1

Table 5 float32 classification report

Label	Precision	Recall	F-1 score
<i>SqueezeNet</i>			
Mask	0.9	1	0.95
No-mask	1	0.89	0.94
<i>Modified squeezeNet</i>			
Mask	0.96	1	0.98
No-mask	1	0.95	0.98
<i>Our model</i>			
Mask	0.98	1	0.99
No-mask	1	0.98	0.99

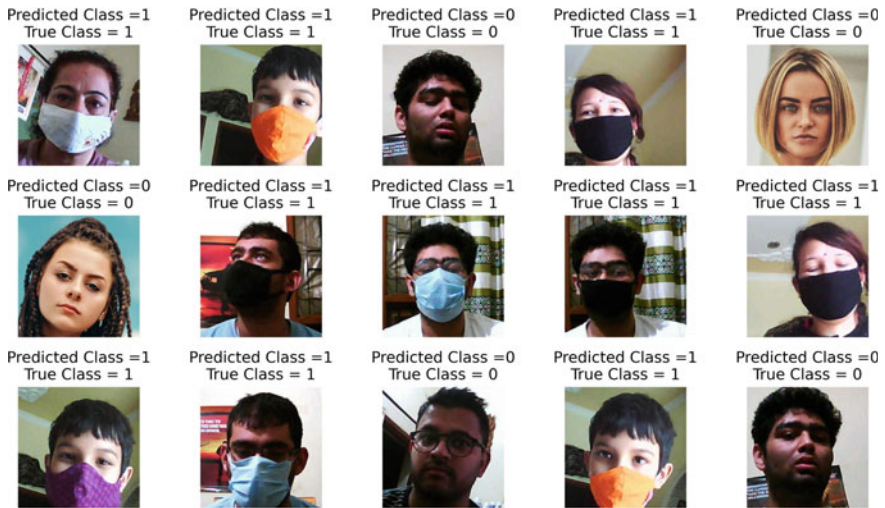


Fig. 6 Some predictions made by our int8 model

In this research, dropout was used after every layer, and it made a significant difference in the test accuracy achieved. Observing the proposed model’s architecture in Fig. 2, it was found that,

“Dropout added after every layer seems to significantly improve the generalization of smaller models.”

Interpolation augmentation, as suggested in [22], was used in the proposed model, improving generalization, and corroborating the statement,

“Interpolation Augmentation seems to improve generalization for resource-constrained endpoints.”

6 Conclusion

In this research work, an extremely small and well-generalizable CNN-based solution has been proposed for face mask recognition on edge devices with extreme resource constraints. The solution has been deployed on a microcontroller development board called OpenMV Cam H7. The model is just 138 KB in size and runs at 30 FPS on the board. It has a test accuracy of 99.83%.

It has been shown that aggressive regularization through dropout might be useful for developing extremely generalizable CNN architectures for problems like binary classification. The method proposed in this paper is universal and applicable to any microcontroller architecture. The methodology used in this research work can also be

used to build and deploy architectures for other challenging problems. The pipeline followed here, which includes, dataset construction, training on float32, quantization to int8, and deployment on edge devices, is applicable to a wide spectrum of resource-constrained, intelligent edge solutions.

Future avenues of research include building systems that are more robust to noise and can work on even smaller microcontrollers. Work can be done on building datasets that include images from more diverse sources. Novel quantization schemes can be developed for converting float32 to int8. Smaller precision can be experimented with including 6-bit, 4-bit and binarized neural networks too.

References

1. A.J. Paul, Recent advances in selective image encryption and its indispensability due to covid-19, in *IEEE Recent Advances in Intelligent Computational Systems (RAICS)* (2020). <https://doi.org/10.1109/RAICSS1191.2020.9332513>
2. Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, in *Intelligent Signal Processing* (IEEE Press, 2001), pp. 306–351
3. A. Chavda, J. Dsouza, S. Badgujar, A. Damani, Multi-stage CNN architecture for face mask detection. arXiv e-prints [arXiv:2009.07627](https://arxiv.org/abs/2009.07627) (2020)
4. M. Loey, G. Manogaran, M. Taha, N. Khalifa, A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the covid-19 pandemic. *Measur. J. Int. Measur. Confederation* **167** (2021). <https://doi.org/10.1016/j.measurement.2020.108288>
5. M. Jiang, X. Fan, H. Yan, RetinaMask: a face mask detector. arXiv e-prints [arXiv:2005.03950](https://arxiv.org/abs/2005.03950) (2020)
6. T. Meenpal, A. Balakrishnan, A. Verma, Facial mask detection using semantic segmentation, in *2019 4th International Conference on Computing, Communications and Security (ICCCS)*, pp. 1–5 (2019). <https://doi.org/10.1109/ICCCS.2019.8888092>
7. M. Loey, G. Manogaran, M.H.N. Taha, N.E.M. Khalifa, Fighting against covid-19: a novel deep learning model based on yolo-v2 with resnet-50 for medical face mask detection. *Sustain. Cities Soc.* **102600** (2020). <https://doi.org/10.1016/j.scs.2020.102600>. <http://www.sciencedirect.com/science/article/pii/S2210670720308179>
8. G. Jignesh Chowdary, N. Singh Punn, S.K. Sonbhadra, S. Agarwal, Face mask detection using transfer learning of inceptionV3. arXiv e-prints [arXiv:2009.08369](https://arxiv.org/abs/2009.08369) (2020)
9. B. Roy, S. Nandy, D. Ghosh, D. Dutta, P. Biswas, T. Das, Moxa: a deep learning based unmanned approach for real-time monitoring of people wearing medical masks. *Trans. Ind. Nat. Acad. Eng.* **5**(3), 509–518 (2020). <https://doi.org/10.1007/s41403-020-00157-z>
10. R. Banner, Y. Nahshan, D. Soudry, Post training 4-bit quantization of convolutional networks for rapid-deployment, in *Advances in Neural Information Processing Systems*, vol. 32. ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox, R. Garnett (Curran Associates, Inc., 2019), pp. 7950–7958. <https://proceedings.neurips.cc/paper/2019/file/c0a62e133894cdce435bcb4a5df1db2d-Paper.pdf>
11. Y. Nahshan, B. Chmiel, C. Baskin, E. Zheltonozhskii, R. Banner, A.M. Bronstein, A. Mendelson, Loss aware post-training quantization. arXiv e-prints [arXiv:1911.07190](https://arxiv.org/abs/1911.07190) (2019)
12. R. Zhao, Y. Hu, J. Dotzel, C. De Sa, Z. Zhang, Improving neural network quantization without retraining using outlier channel splitting, in *Proceedings of the 36th International Conference on Machine Learning, Proceedings of Machine Learning Research*, vol. 97 ed. by K. Chaudhuri, R. Salakhutdinov (eds.) (PMLR, Long Beach, California, USA 2019), pp. 7543–7552 <http://proceedings.mlr.press/v97/zhao19c.html>

13. Y. Choukroun, E. Kravchik, F. Yang, P. Kisilev, Low-bit quantization of neural networks for efficient inference, in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pp. 3009–3018 (2019). <https://doi.org/10.1109/ICCVW.2019.00363>
14. B. Jacob, S. Kligys, B. Chen, M. Zhu, M. Tang, A. Howard, H. Adam, D. Kalenichenko, Quantization and training of neural networks for efficient integer-arithmetic-only inference, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2018)
15. R. Gong, X. Liu, S. Jiang, T. Li, P. Hu, J. Lin, F. Yu, J. Yan, Differentiable soft quantization: bridging full-precision and low-bit neural networks, in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2019)
16. S. Ge, Z. Luo, S. Zhao, X. Jin, X. Zhang, Compressing deep neural networks for efficient visual inference, in *2017 IEEE International Conference on Multimedia and Expo (ICME)* (2017), pp. 667–672. <https://doi.org/10.1109/ICME.2017.8019465>
17. Z. Dong, Z. Yao, A. Gholami, M.W. Mahoney, K. Keutzer, Hawq: Hessian aware quantization of neural networks with mixed-precision, in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2019)
18. Z. Dong, Z. Yao, Y. Cai, D. Arfeen, A. Gholami, M.W. Mahoney, K. Keutzer, Hawq-v2: Hessian aware trace-weighted quantization of neural networks, in *Advances in Neural Information Processing Systems 33 Pre-proceedings (NeurIPS 2020)* (2019)
19. Z. Yao, Z. Dong, Z. Zheng, A. Gholami, J. Yu, E. Tan, L. Wang, Q. Huang, Y. Wang, M.W. Mahoney, K. Keutzer, HAWQV3: dyadic neural network quantization. arXiv e-prints [arXiv:2011.10680](https://arxiv.org/abs/2011.10680) (2020)
20. I. Abdelkader, Y. El-Sonbaty, M. El-Habrouk, Openmv: a python powered, extensible machine vision camera. arXiv e-prints [arXiv:1711.10464](https://arxiv.org/abs/1711.10464) (2017)
21. 32-bit Arm® Cortex®-M7 480MHz MCUs, up to 2MB Flash, up to 1MB RAM, 46 com. and analog interfaces (2019). <https://www.st.com/resource/en/datasheet/stm32h743vi.pdf>
22. A.J. Paul, P. Mohan, S. Sehgal, Rethinking generalization in american sign language prediction for edge devices with extremely low memory footprint, in *IEEE Recent Advances in Intelligent Computational Systems (RAICS)* (2020). <https://doi.org/10.1109/RAICSS51191.2020.9332480>
23. A. Jangra, Face mask 12k images dataset (2020). <https://www.kaggle.com/ashishjangra27/face-mask-12k-images-dataset>
24. D. Makwana, Face mask classification (2020). <https://www.kaggle.com/dhruvmak/face-mask-detection>
25. F.N. Iandola, S. Han, M.W. Moskewicz, K. Ashraf, W.J. Dally, K. Keutzer, SqueezeNet: alexNet-level accuracy with 50x fewer parameters and <0.5MB model size. arXiv e-prints [arXiv:1602.07360](https://arxiv.org/abs/1602.07360) (2016)

Analysis of Darknet Traffic for Criminal Activities Detection Using TF-IDF and Light Gradient Boosted Machine Learning Algorithm



Romil Rawat, Vinod Mahor, Sachin Chirgaiya, Rabindra Nath Shaw, and Ankush Ghosh

Abstract Darkweb also called as sinkholes, blackholes, network telescopes, and darknet is the environment and the most favorable platform for illegal activities due to hidden IP address and therefore counted as unused address space, which is not available for normal user, and the anonymous behavior acts as catalyst for criminal or unauthorized behavior conduction. It is very difficult to suddenly trace the location of malicious activity origin but by traffic analysis and understanding the patterns, suspicious activities including email communication, audio–video streaming, chatting P2P, browsing data, chatting, and voice over Internet protocol constitute the hidden world web traffic. Several methods have been deployed to analysis and classify darkweb network traffic. The proposed work detects worms, dos attack, backdoor, DDos attack, RDoS attack, spam, and malicious contents. In the proposed work, term frequency-inverse document frequency (TF-IDF) and light gradient boosted machine algorithm method has been implemented on darknet traffic data. The light gradient boosted machine algorithm shows the value of 98.97% as accuracy and thus outperforms the other algorithms based on experiment values.

Keywords Darknet traffic · TF-IDF · Light gradient boosted machine learning · Cyberattack · Data mining

R. Rawat · S. Chirgaiya

Department of Computer Science Engineering, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore, India

V. Mahor

Department of Computer Science Engineering, Gwalior Engineering College, Gwalior, India

R. N. Shaw

Department of Electrical, Electronics and Communication Engineering, Galgotias University, Greater Noida, India

e-mail: r.n.s@ieee.org

A. Ghosh (✉)

School of Engineering and Applied Sciences, The Neotia University, Sarisha, West Bengal, India

1 Introduction

The denial of service (DoS) attack or request of unstructured packet flood generated from unauthorized host in a short or continues period of span to consume software and hardware resources for blocking the legitimate service like malicious program checker and system security routine checker and thus opens port for vulnerabilities [1], and they are TCP SYN flooding attack, TCP attack/synchronize (TCP SYN), Internet control message protocol (ICMP), [2] user datagram protocol (UDP).

Important Features:

- Unused IP address constitutes darkweb space [3], because billions of IP address are there in IPv4, and not all are used for system allocations and traced for their behavior.
- The unused IP address is searched and thus used by cybercriminals for generating [4] unauthorized traffic in the form of malware, ransomware, and Trojan horse.
- This unused IP address is the hub of flooding attack (backscatter, stealth scan, socket, port, scan, DoS, DDoS, and RDoS) [5] and even does not leave log on victim computer once the system is compromised.
- This address even always scan open ports, connections, host so that backdoor attack can be triggered.
- Most and easy attack procedure followed by criminals [6] is to use phishing text messages, mail message, image, post, and audio–video link to secretly get login credentials by designing the forged link.
- Other attacks created by cyber attackers to tricks users are virus, spam, fraud, or identity [7] theft.

Network traffic dump [8] is presented in Fig. 1. It contains useful information of packet, route, and with confidential data analysis.

Therefore, the objective of the proposed work is to analyze criminal activities detection using TF-IDF and light gradient boosted machine algorithm of darknet traffic platform.

Fig. 1 Network traffic dump

Name	Description
Source Port	Source port-transport layer
TCP Flags	TCP flags
Transport Protocol	Protocol integer code-transport layer
Packet Direction	Packet direction
Packet Type	normal or attack
TimeStamp	Received packet time
PacketSize	Actual length of packet in byte
SourceIP	Source IP address -network layer
Destination IP	Destination IP address-network layer
Destination Port	transport layer destination port

The rest of the paper is organized as follows. Section 2 covers the literature review. The methodology and propose algorithm on how the experiment will be carried out are presented in Sect. 3, Sect. 4 describes the dataset structures, and lastly, Sect. 5 presents the results and discussion on the finding. Finally, Sect. 6 concludes this paper.

2 Related Work

Farhad Soleimani et al. presented K-means and fuzzy K-means data mining clustering approaches for intrusion detection framework. They gathered dataset from KDD cup which has 41 traits. They utilize k-means and fuzzy k-means ways to deal with recognize the kind of DoS vulnerability. They reasoned that the fuzzy k-means technique accomplished somewhat in a way that is better than k-means strategy in distinguishing flooding attack [7].

Han et al. [6] utilized data mining technique for intrusion detection. They gathered dataset from MIB network diverse time span. The creators applied data entropy to choose most critical properties from the whole set of dataset. They presumed that the methodology has given great execution to choose better element credits and high accuracy detection rate.

Mill operator, W. Deritrik, W. Hu et al. proposed recurrence histogram approaches for identifying strange bundles dependent on data stream mining. Network data is gathered from various databases DARPA and MCPAD datasets. The Den Stream approach treats singular parcels as focuses and are hailed as typical or noxious dependent on whether these focuses are ordinary and exceptions. They used a histogram way to deal with fabricate the histogram model for new parcel payload. They utilized Pearson relationship for figuring between two histograms. From the outcome, they presumed that histogram-based detection calculation accomplished minimal better execution yet required a larger number of quantities of highlights than the clustering-based algorithm [9].

Ghanshyam Prasad Dubey et al. proposed RST and incremental SVM ways to deal with cyber illicit behavior originated by intruder. The incremental SVM approach enhanced the expanded identification for intrusion detection. They noticed that the RST and addition SVM approaches are compelling to diminish the space thickness of data [10].

Shailendra Kumar et al. presented harsh set hypothesis and support vector machine for dimensionality reduction in intrusion detection. They tried different things with KDD cup dataset. They applied harsh set hypothesis to choose the main credits from KDD cup dataset. A near investigation between SVM with unique 41 datasets and reduction dataset is introduced. They presumed that the proposed calculation is entirely solid for intrusion detection [11].

Heba F. Eid et al. analyzed intrusion detection framework utilizing support vector machines with principal component analysis draws near. They tried their proposed model on NSL-KDD dataset. The PCA approach is used to diminish the quantity

of highlights to diminish the unpredictability of the framework. Their outcomes demonstrated that the proposed framework is able to accelerate the cycle of intrusion detection and to lessen the memory space and CPU time cost [12].

S. Mehibs and S. Hashim proposed a network intrusion detection framework (NIDF) based on fuzzy c_mean calculation in distributed computing surroundings [13]. They propose a network intrusion detection framework in a cloud computing environment dependent on backpropagation neural network (BPNN) [14].

Bhaya, W. and, Ebadymanaa, amalgamated unsupervised data mining techniques (UDMT) as intrusion detection framework. The entropy idea in term of windowing the approaching IP data packets is implemented with data mining technique utilizing clustering using representative as cluster investigation to recognize the DDoS vulnerability in network stream [15].

Vivek K. Kshirsagar et al. proposed decision tree (DT) techniques for exploring and assessing intrusion detection. The creators gathered dataset from DARPA dataset which has an alternate kind of cybercrime records and patterns for analyzing cyberattacks. The SVM gave preferable outcome over DT on DoS vulnerability class. They saw that the DT is more proficient for identifying the intrusion [16].

3 Proposed Work

The objective of proposed work is to analyze darkweb traffic for vulnerabilities, unauthorized behavior, malicious message, cybercriminal crowd identification, and packets tracing in network data stream.

3.1 Design Architecture

Figure 2 shows the process of darknet traffic data collected from standard data source and, then cleaning, preprocessing is done using data mining techniques, and TF-IDF is applied with light gradient boosted machine learning algorithm to generate the final behavior of processed data.

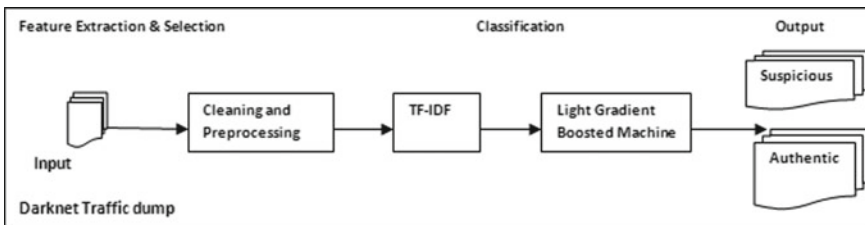


Fig. 2 Design architecture for analyzing network traffic dump

3.2 TOR Traffic Data Classification

Figure 3 classifies the data generated from the onion router (TOR) [17]. The data travels from TOR browser using key and hash value generated by TOR directory with relay nodes and passes across darkweb router and thus makes it identity hidden and anonymous. The traffic data is analyzed, and useful information and patterns are analyzed for further design development to trace host generating darkweb crime and malicious behavior and secret operating illicit [18] markets of drug trafficking, money laundering, organ trafficking, cyberterrorism, radicalization and recruitment [19], contract crime induction, illicit gambling and weapon business, cyber extortion, human trafficking and abuse, and selling of stolen antique items [20].

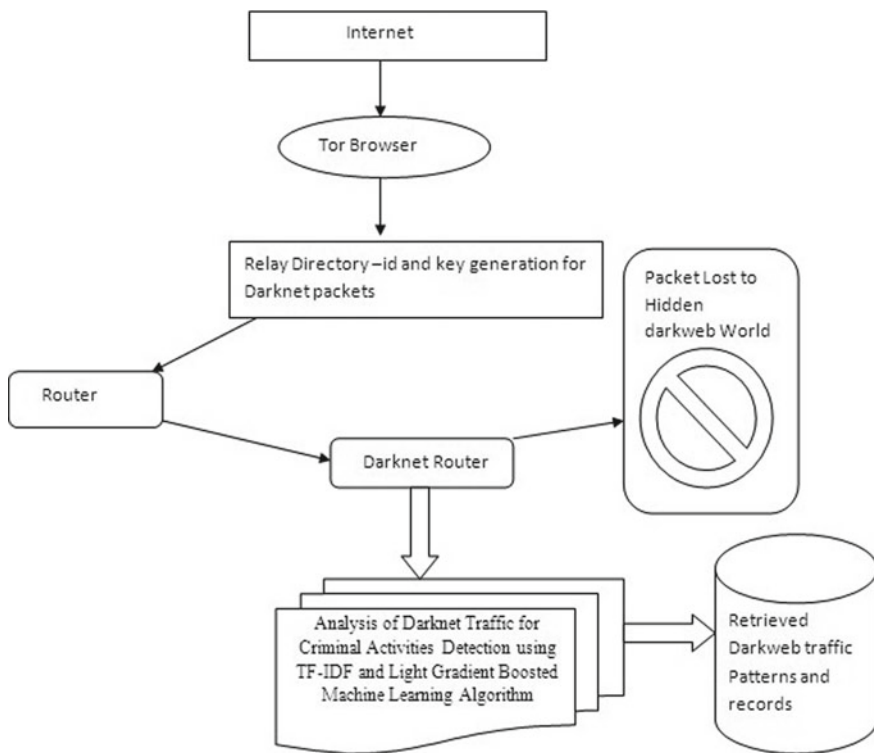


Fig. 3 Classification of data generated from the onion router (TOR)

4 Dataset

Public dataset is collected from different sources listed in Fig. 4: designed for darkweb traffic analysis for proposed research.

The phases involve preprocessing, feature extraction, selection, and classification.

4.1 Preprocessing

The data preprocessing is the primary stage in which the unstructured data are transferred in proper form of structured data (Fig. 5).

4.2 Feature Extraction and Selection

TF-IDF [23] is the frequently weighting technique used to in the vector space model document and text mining for analyzing the features. It is a factual strategy to quantify the significant of a word in the document to the entire corpus. The term frequency is just determined in relation to the quantity of events a word that shows up in the document and typically standardized in certain quadrant somewhere in the range of 0 and 1 to take out inclination toward protracted documents [11]. To develop the record of terms in TF-IDF, accentuation is eliminated, and all text are lowercase during tokenization. The initial two-letter TF or term frequency alludes to how significant on the off chance that it happens all the more oftentimes in a document. The higher TF reflects to the more assessed that the term is huge in particular documents. Moreover, IDF or inverse document frequency is determined on how inconsistent a word or term is in the documents [11]. The weighted worth is assessed utilizing the entire preparing dataset. The possibility of IDF is that a word is not viewed as acceptable contender to speak to the document on the off chance that it is happening habitually in the entire dataset as it very well may be the stop words or basic words that are conventional.

Dataset Name	Information
DUTA – Darknet usage text addresses	Suspicious Domains
ISCXTor2016	Tor-nonTor dataset
ISCXVPN2016	VPN-nonVPN dataset
Dark Net Marketplace Data (Agora 2014-2015)	listings of drugs, weapons
UCSD Network Telescope -- Darknet Scanners Dataset	Network Dump
NUST2009	Malware Data
UNSW-NB2015	Malware Data
NICT /16 darknet sensor	TCP SYN packets

Fig. 4 Dataset of darkweb traffic [21, 22]

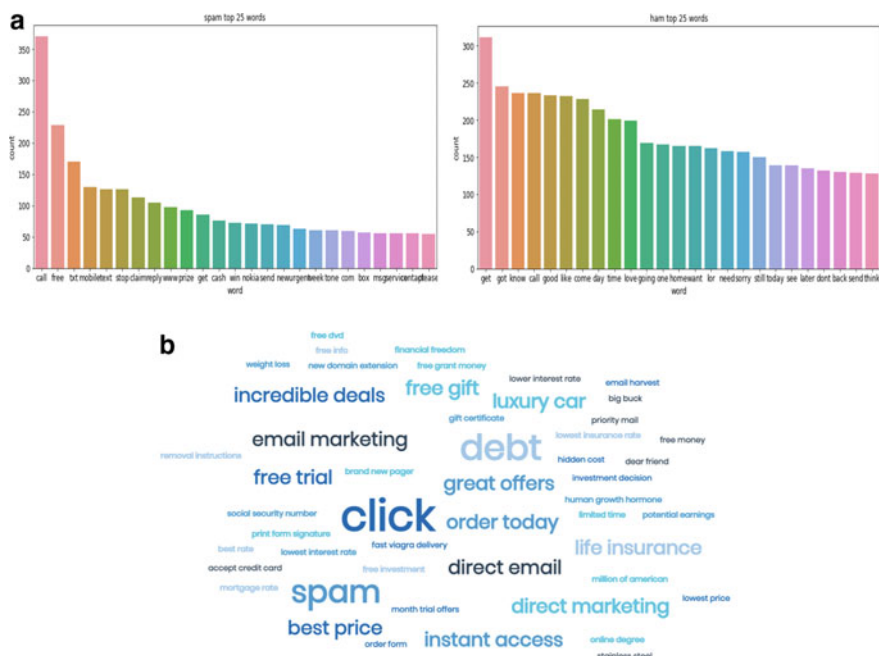


Fig. 5 **a** Classification of frequent words used in cyber communication, and **b** word cloud of cyber words mostly used for enticing the user to engage in malicious activity

Thus, just rare words interestingly of the whole dataset are significant for those documents. TF-IDF does not just survey the significance of words in the documents; however, it likewise assesses the significance of words in document information base or corpus. In this sense, the word frequency in the document will expand the weight of words relatively, however, will at that point be counterbalanced by corpus' promise frequency [8]. This critical quality of TF-IDF expects that there are a few words that show up more frequently contrasted with others documents. Figure 6 described the equation for documents analysis.

4.3 Light Gradient Boosted Machine Algorithm (LightGBM)

LightGBM, shortly defined for light gradient boosted machine [8], is a gradient boosting framework for machine dependent on decision tree calculations and utilized for ranking, characterization, and other machine learning algorithms. The improvement center is around execution and adaptability. The framework underpins various calculations including GBM—gradient boosting machine, MART—multiple additive regression trees, GBT—gradient boosting tree, GBDT—gradient boosted

Fig. 6 TF-IDF equation and calculations [24]

$$TFIDF \text{ score for term } i \text{ in document } j = \frac{TF(i, j) * IDF(i)}{\text{where}}$$

IDF = Inverse Document Frequency
TF = Term Frequency

$$TF(i, j) = \frac{\text{Term } i \text{ frequency in document } j}{\text{Total words in document } j}$$

$$IDF(i) = \log_2 \left(\frac{\text{Total documents}}{\text{documents with term } i} \right)$$

and
t = Term
j = Document

decision tree, GBRT—gradient [25] boosted regression trees, and RF—random forests.

It is intended to be appropriated and proficient with the accompanying featured points:

- Faster preparing speed and higher productivity.
- Lower memory utilization.
- Better exactness.
- Support of equal and GPU learning.
- Capable of taking care of enormous scope information (Fig. 7).

Algorithm 1: Gradient_TreeBoost	
1	$F_0(\mathbf{x}) = \arg \min_{\gamma} \sum_{i=1}^N \Psi(y_i, \gamma)$
2	For $m = 1$ to M do:
3	$\tilde{y}_{im} = - \left[\frac{\partial \Psi(y_i, F(\mathbf{x}_i))}{\partial F(\mathbf{x}_i)} \right]_{F(\mathbf{x})=F_{m-1}(\mathbf{x})}, i = 1, N$
4	$\{R_{lm}\}_1^L = L$ - terminal node $tree(\{\tilde{y}_{im}, \mathbf{x}_i\}_1^N)$
5	$\gamma_{lm} = \arg \min_{\gamma} \sum_{\mathbf{x}_i \in R_{lm}} \Psi(y_i, F_{m-1}(\mathbf{x}_i) + \gamma)$
6	$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \nu \cdot \gamma_{lm} 1(\mathbf{x} \in R_{lm})$
7	endFor

Fig. 7 LightGBM pseudo code [21]

5 Result and Outcomes

In the proposed work, performance parameters precision, false alarm value, detection rate value, accuracy, and *F*-measure or *F*-score are evaluated for the comparatively study of algorithms. The work is focused on dataset—UNSW-NB15, to continue and compare the result values in a same taken environment chain. Figure 8 given below provides the details about comparative study of proposed and available algorithm parameters.

5.1 Parameter Comparison

Algorithms are compared using different parameters to check the performance.

5.2 Accuracy Diagram

Figure 9 systematically outlines the values of different algorithms used in research. The light gradient boosted machine algorithm shows the value of 98.97% as accuracy and thus outperforms the other algorithms based on experiment values.

The TF-IDF + LightGBM accomplish the best performance and beat the other assessed algorithms in this investigation regarding accuracy ratio. It precisely characterizes 98.97% while detection rate—100%, false alarm rate—0.01, and *F*-score value—83.73%. The other algorithms are named as statistical model (BMM-2017), statistical model (GAA-2017), TF.IDF_BFR 2019 achieved the accuracy 93.4%, 91.8%, and 98.76%, respectively. Several other parameters are also compared and mentioned in Fig. 8.

No.	Method	Accuracy	Detection Rate	False Alarm Rate	F-Measure or F-Score	Dataset
1	Statistical model(BMM) 2017	93.4%	92.7%	5.9%	--	UNSW-NB15
2	Statistical model GAA-ADS 2017	91.8%	91.0 %	5.8%	--	UNSW-NB15
3	TF.IDF_BFR 2019	98.76%	100%	0.02%	79.28%	UNSW-NB15
4	Proposed work(TF-IDF+ Gradient Boosting Classification)	98.97%	100%	0.01	83.73%	UNSW-NB15

Fig. 8 Comparative study of proposed and available algorithm parameters

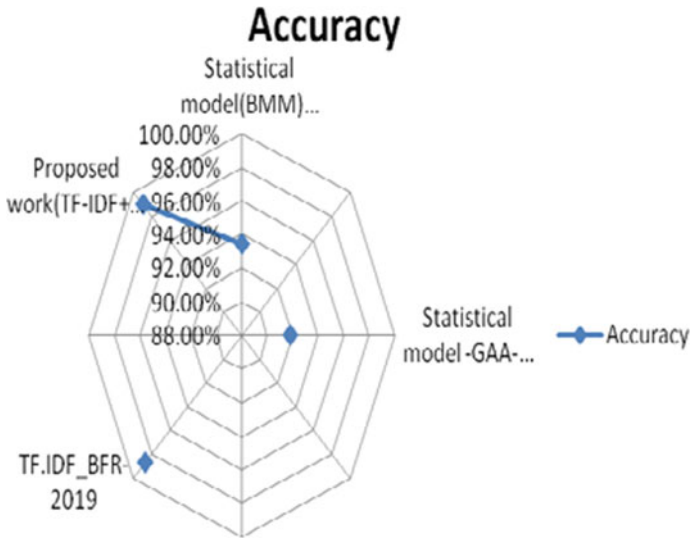


Fig. 9 Accuracy diagram

6 Conclusion

The proposed work detects and analyzes suspicious information's and packet data found in network traffic for understanding the behavior of darkweb channels. The data mining and machine learning are used with standard dataset to analyze and detect various types cyber threats creating nuisance behavior in network. The proposed work is not deployed previously and thus used here.

7 Future Work

Intelligent live tracing algorithm for traffic tracing, analyzing, and blocking with detailed report should be developed, so that cybercriminal's activities could be blocked with immediate action and plan.

References

1. N. Hashimoto, S. Ozawa, T. Ban, J. Nakazato, J. Shimamura, A darknet traffic analysis for IoT malwares using association rule learning, in *INNS Conference on Big Data and Deep Learning 2018*. Procedia Comput. Sci. **144**, 118–123 (2018)
2. T. Ban, M. Eto, S. Guo, D. Inoue, K. Nakao, R. Huang, A study on association rule mining of darknet big data, in *Proceedings of International Joint Conference on Neural Networks* (2015), pp. 1–7
3. C. Stocker, J. Horchert, Mapping the internet: A hacker’s secret internet census. Spiegel Online, March 22, 2013
4. E.L. Malecot, D. Inoue, The Carna botnet through the lens of a network telescope, in *Foundations and Practice of Security, LNCS*, vol. 8352, ed. by J. Danger et al. (Springer, Berlin, 2014), pp. 426–441
5. R. Agrawal, T. Imielinski, A. Swami, Mining association rules between sets of items in large databases. ACM SIGMOD Rec. **22**(2), 207–216 (1993)
6. J. Han, J. Pei, Y. Yin, Mining frequent patterns without candidate generation. ACM SIGMOD Rec. **29**(2), 1–12 (2000)
7. X. Du, Y. Yang, X. Kang, *Research of Applying Information Entropy and Clustering Techniques Network Traffic Analysis* (IEEE, 2008), 978-0-7695-3508-1
8. <https://www.kaggle.com/philipjames11/dark-net-marketplace-drug-data-20142015>
9. F.S. Gharehchopogh, N. Jabbari, Z.G. Azar, Evaluation of fuzzy k-means and k-means clustering algorithms in intrusion detection systems. Int. J. Sci. Technol. Res. **1**(11), 66–71 (2012)
10. Z. Miller, W. Deitrick, W. Hu, Anomalous network packet detection using data stream mining. J. Inf. Secur. **2**(4), 158–168 (2011)
11. G.P. Dubey, N. Gupta, R.K. Bhujade, A novel approach to intrusion detection system using rough set theory and incremental svm. Int. J. Soft Comput. Eng. (IJSCE) **1**, 663–667 (2011)
12. R.-C. Chen, K.-F. Cheng, Y.-H. Chen, C.-F. Hsieh, Using rough set and support vector machine for network intrusion detection system, in *2009, First Asian Conference on Intelligent Information and Database Systems* (2009), pp. 465–470
13. V.K. Kshirsagar, S.M. Tidke, S. Vishnu, Intrusion detection system using genetic algorithm and data mining: an overview. Int. J. Comput. Sci. Inf. **1**(4) (2012). ISSN (PRINT): 2231-5292
14. S. Mehibs, S. Hashim, Proposed network intrusion detection system based on fuzzy c_mean algorithm in cloud computing environment. JUBPAS **26**(2), 27–35 (2017)
15. S. Mehibs, S. Hashim, Proposed network intrusion detection system in cloud environment based on back propagation neural network. JUBPAS **26**(1), 29–40 (2017)
16. H.F. Eid, A. Darwish, E. Hassanien, A. Abraham, Principle components analysis and support vector machine based intrusion detection system, in *10th International Conference on Intelligent Systems Design and Applications (ISDA)* (IEEE, 2010), pp. 363–367
17. W. Bhaya, M. Ebadymanaa, DDoS attack detection approach using an efficient cluster analysis in large data scale, in *2017 Annual Conference on New Trends in Information and Communications Technology Applications, NTICT 2017* (2017)
18. <https://gvis.unileon.es/dataset/duta-darknet-usage-text-addresses-10k/>
19. <https://www.unb.ca/cic/datasets/darknet2020.html>
20. <https://www.unb.ca/cic/datasets/tor.html>
21. <https://mungingdata.wordpress.com/2017/11/25/episode-1-using-tf-idf-to-identify-the-signal-from-the-noise/>
22. <https://www.sciencedirect.com/science/article/abs/pii/S0968090X15000741>
23. <https://www.unb.ca/cic/datasets/vpn.html>
24. https://www.researchgate.net/publication/330523899_Anomaly_Detection_System_for_Internet_Traffic_based_on_TF-IDF_and_BFR_Clustering_Algorithms
25. https://www.caida.org/data/passive/telescope-darknet-scanners_dataset.xml

Risk Detection in Wireless Body Sensor Networks for Health Monitoring Using Hybrid Deep Learning



Anand Singh Rajawat, Kanishk Barhanpurkar, Rabindra Nath Shaw, and Ankush Ghosh

Abstract The Internet of Things (IoT) idea has arisen as interconnected components of the healthcare tracking facilities of smart linked healthcare networks. Hard sensor-based data aggregation with the help of devices in the form of wearables or intrusive samples attached with the acquisition of soft sensors like crowd sensor results in the aggregated sensor data being concealed in patterns. This problem is tackled through several secret stages of interpretation of deep learning techniques. In this research work, we proposed hybrid deep learning (HDL) techniques to develop estimation and enhance quality of smart health services on health monitoring data. We also showed a detailed comparison of methods on the basis of health surveillance types. Hence, our proposed models work for risk detection in health information which will help us to increase the efficiency of existing healthcare industry.

Keywords WBAN · Back propagation neural network · Deep learning

1 Introduction

Health surveillance systems are based on information and communication technologies primarily to connect people and infrastructure in order to enhance the quality of life in general. Efficient utilization and security of energy, industrial stability and

A. S. Rajawat (✉)

Department of Computer Science Engineering, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore, India

K. Barhanpurkar

Department of Computer Science and Engineering, Sambhram Institute of Technology, Bengaluru, Karnataka, India

R. N. Shaw

Department of Electrical, Electronics and Communication Engineering, Galgotias University, Greater Noida, India

e-mail: r.n.s@ieee.org

A. Ghosh

School of Engineering and Applied Sciences, The Neotia University, Sarisha, West Bengal, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

683

S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,

Lecture Notes in Electrical Engineering 756,

https://doi.org/10.1007/978-981-16-0749-3_54

sustainable development support the basic criteria of health monitoring systems, among other aspects, while in the quick accessibility and access to reliable medical care, those targets are accomplished by adequately mandating and transmitting data from dedicated networks (Zhao et al., 2019). Wireless body area networks (WBAN) are formed with miniature sensors with wireless communication capabilities and have been developed by scientists and researcher's capabilities that may be inserted in the patient's body. The instruments continuously collect biological signals with the help of different sensors and relay captured data to the local analysis unit (LPU) from patients in the hospitals and quarantine centers (Xia et al., 2020). LPUs should be durable and capable of processing health data in real time which increase health warnings for healthcare providers to rapidly respond by taking proactive steps when they notice the declining health status of patients or any emergency situation. For preservation and long-term analysis, LPU data may also be distributed to remote databases (DB). WBANs have several benefits including allowing physicians to monitor patient-specific attributes irrespective of locality, enhancing diagnostic precision and effectiveness and decreasing the healthcare cost per patient by providing a suitable medium for physicians to constantly monitor patient health. In this paper, the usefulness of hybrid deep learning model for health data classification through hybrid profound learning is increased. WBAN sensor nodes are vulnerable to both hardware and software problems such as defective modules, sensor activation, battery depletion and dislocation (Ding et al., 2018). The sensor readings are both untrustworthy and unreliable due to the restricted hardware resources such as decreased computing power, insufficient storage and memory and transmission. Data processing and delivery by individual sensors are often likely to require many forms of irregularity, such as interference, noise, misplacements by sensors, sweater patients, inadequate power resources and external hacks. This could lead to unforeseen outcomes, inaccurate warnings and evaluation and a drop of public confidence in these programs (Jones et al. 2010). High false alarm rate and inaccurate measurements thus directly affect the public reputation of WBANs in particular where reliability, as it does in the medical field, is extremely important. For example, if the sensor is wrongly connected with a pulse oximeter or external fluorescent light irradiates to the infrared sensor, it can result in mistaken metrics. Number of authors noticed that, compared with some other concerns (i.e., network malfunction, data transfer), the first cause of unreliability in medical WSNs is sensing components. The transmission of false data by nodes has an adverse impact on the accuracy of the collected data, which could affect the diagnosis of patients. This, in turn, will lead to life-threatening circumstances in which people can receive emergency warnings based on node defects of blue code (Alameen et al., 2020). As a result, it is highly important for incorrect node level steps to be observed and for false alerts to be reduced to discriminate between patient irregularities and node faults. The remainder was grouped in the following manner. In Sect. 2, we discuss similar studies on identification of abnormalities and learning devices for WSN medical applications. Section 3 quickly explores our identification method for WSN health data classification and 3.1 and 3.2 big health data classification using hybrid-based learning. We present

our test results, which analyze the solution proposed with real patient data. In this Section 4, we concludes the paper with a discussion of future work results and plans.

2 Related Work

The pathological tests are made for patient irregularities and node defects, both of which must be identified with the greatest possible precision. This can only be achieved through a mechanism for the medical data classification to detect and extract abnormal patterns and correlations in the information and to separate ill people from faulty sensor systems (Liu et al., 2020). Anomalies are classified however as variations from a standard model that is dynamically modified from the sensed data. In most sensor node hardware, external sources are usually needed to use these distributed approaches. As a consequence, their precision is less than clustered methods that use a global picture for spatial–temporal analysis. Correlations between physical dimensions that occur in time and space must therefore be manipulated to identify and retrieve abnormal quantities to guarantee effective operations and precise diagnosis (Chen et al., 2020). In general, the tracked characteristics for inaccurate measurements have no spatial or temporal association. The key emphasis of this paper is the creation of a HDL model for medical data classification. However, physiological characteristics are closely associated, and the heart rate and respiration levels are elevated at the same time, usually in at least two or three conditions, e.g., in atrial fibrillation (AF) and asthma (Hussain et al., 2020). Our approach would boost the efficiency of WBANs for patient monitoring. The key task is to detect and remove WBAN data irregularities and to discriminate between unusual, vital patient signs and inadequate sensor measurements (Noseworthy et al., 2020). Furthermore, the number of author was attempting to mitigate false alerts caused by irregular data sensors. In smart health apps, we include a brief analysis of deep learning methods that are applicable for sensor data (Jiang et al., 2019). This research is focused on the fact that the main enablers of digital health technologies are deep learning techniques. This is because the traditional computer teaching techniques in high-dimensional data have efficiency and precision problems (Shankar et al., 2020). Deep learning is not necessarily a complete alternative for machine learning, but an important method in many fields, such as smart health, to deal with dimensional problems. To this end, the goal of this paper is to highlight hybrid’s development of deep learning strategies in intelligent health while simultaneously presenting potential directions by exploring still important problems and open issues.

3 Proposed Methodology

The implementation of the risk detection in wireless body sensor networks for health monitoring WBAN data analysis model has established deep learning as a feasible

Table 1 Comparative analysis of studies carried out on different healthcare applications and deep learning methods

Study Reference	Year Focus	Paper Type	Number of primary studies	Problem	Application	Deep Learning Technique
He L. et al.	2018	Journal	16	Depression detection	Predictive analysis	Convolutional neural network (CNN)
Yang, L. et al.	2017	Conference	29	Depression detection	Predictive analysis	Deep belief networks (DBN)
Cen L. et al.	2019	Conference	3	Sleep apnea detection	Predictive analysis	Convolutional neural network (CNN)
Yamanakkan avar, N. et al.	2020	Journal	N/A	Brain tissue classification	Medical image processing	Artificial neural network (ANN)
Jo T. et al.	2019	Journal	8	Alzheimer's disease	Medical image processing	Convolutional neural network (CNN)
Siar M. et al.	2019	Conference	2	Brain tumor detection	Medical image processing	Convolutional neural network (CNN) + softmax layer
Ahmed O. et al.	2019	Conference	9	Gene expression classification	Bioinformatics	Convolutional neural network (CNN) and recurrent neural network
						neural network (RNN) and deep neural network (DNN).
Nazari I. et al.	2018	Journal	141	Protein RNA and DNA slicing	Bioinformatics	Artificial neural network (ANN)

solution to reduce network capital and operating costs and boost the network's lives. Currently, most deep learning techniques use a mixture of the extraction of characteristics and modality specific algorithms used risk detection in WBAN health monitoring data using hybrid deep learning. In order to perform massive background operations, this typically involves a high volume dataset and strong computational tools. In the face of contradictions and complexity in the dataset, these are often still expected to exist, through exhaustive attempts. One of the key benefits of deep learning is function training, where a computer is training on certain datasets and its output reflects the original feature in a useful way most of the time. The uses of sensory algorithms for deep learning are common such as indoor telemedicine. Conventional machinery learning has, however, certain drawbacks such as the failure to maximize non-differentiable interruptive loss functions or the failure to achieve outcomes at all times after a practicable training period. The challenges of deep learning, along with many others, paved the way for deeper learning as a robust learning instrument. In the subsections 3.1, 3.2 and 3.3, we have explained how big data technologies can be used to classify data and deep learning techniques that will be useful in development of model.

3.1 Big Health Data Classification Using Hybrid-Based Learning

A set of algorithmic processes that “imitate the brain” are described in deep learning. Deep learning also involves learning in algorithmic layers. These strata allow hierarchical knowledge to be defined on the basis of basic knowledge. There have been many efforts to build and grow machines that are able to think. Until recently, this effort has become a “top-down,” rule-based learning approach which calls for guidance to be implemented in all circumstances. However, where the rules are minimal, this technique is limited to the number of rules (Ali et al., 2020).

These challenges can be addressed by using the bottom-up approach instead of by learning from rules-based practice. The marked details form the perception. The details on which these are marked were used as an instruction manual for the process of preliminary part (Li et al., 2019). The technique of gaining information is sufficient for applications such as health monitoring. On the other hand, the majority of the data collected from risk detection in WBAN health monitoring data sensors is not appropriately labeled (e.g., maximum heart rate, ECG signals, etc.).

Real-world issues with the processing of multimedia sensor data like voice are problematic because of the potentially unlimited problem area to represent digitally. Thus, it is particularly the existence of multi-dimensional functionalities that sufficiently describes the problem, resulting in an increase in space volume in a way which makes data accessible sparse, while sparse data training does not lead to meaningful results. However, in the treatment of sensory data that are often derived from multimedia sensors, problems of “infinite preference” are prevalent. This gives way to

profound learning as profound learning algorithms severe and/or logical questions, with little to no control on high-dimensional characteristics, must be focused upon (Rathore et al., 2019). The lack of a rule collection renders the device prepared to contend with unpredictable situations. The discovery of complex structures in large datasets is also another aspect of deep learning. For this to happen, profound learning requires a back propagation algorithm to change the internal parameters of each layer depending on the parameters of the previous layer. The representation learning can be reported on partially labeled or unlabeled sensory data, using deep learning (specifically CNN) approaches in coronary artery dysfunction diagnosis and identification from signals obtained by electrocardiograms (ECG). In addition, deep epileptic learning was proposed with edge computation to find epileptogenicity using electromagnetic data in the sense of the emotional well-being which is a crucial condition in human lives to diagnose epileptic disorders at an early level. In this way, the authors use deep belief networks (DBN) and EEG evidence in graded positive and negative emotions.

3.2 Big health data classification

Sensors are key unlocks of the modern IoT networks' artifacts (things). The combination of sensors constitutes a network which, for inferential purposes, aims, among other things, to produce and aggregate data. Until performing some analysis process, the data from sensors must be precisely calibrated. This has contributed to new ways in which sensitive data from sensory and actuator networks can be accurately and appropriately handled. These methods rely on the form of sensed data and their implementations. In sensor and actuator network applications, deep knowledge can be used to effectively process sensory data. For decision-making, the output of a deep learning network may be used in the use of a convolutionary neural network. The objective is to study the factors contributing in risk detection in wireless body sensor networks using hybrid deep learning (Dash et al., 2019). The use of sensors in health care has achieved significant results, and profound research is being used to use sensors and actuators for the proper provision of medical care. In the assessment of Parkinson's level, maximum heart rate, ECG signals, for example, the use of CNNs-DBN is made of data generated by wearable sensors to classify and detect key features of Parkinson's disease of the key features in Parkinson's disease based on data generated from wearable sensors. In contrast to other approaches, the findings of the study revealed that profound teaching strategies worked well with sensors. In addition, in-depth learning techniques, in particular CNNs-DBN, were used to estimate energy use, while wearable health condition monitoring sensors were used (Khan et al. 2020). In addition, a deep learning classification was built for risk detection particularly among the elderly by the use of sensory data from an infrared distance sensor. Besides this, in terms of security, a more robust approach to spoofing and security detection in digital systems has been implemented by integrating biometric sensors and CNNs-DBN.

3.3 *Back Propagation Neural Network (BPNN) for WBAN*

First, sequence-to-sequence is a problem setting. BPNN for WBAN are models, that is, we have to be trained using some data and then be used to predict on new data. A feed-forward neural network looks something like this: Only feeds the input in the forward direction, that is, the output from one unit is only fed to units further ahead. On the other hand, hybrid deep learning looks something like the following.

It is the unit that takes an input layer 0, outputs or hidden layer 3 vector h , and output layer 4 also has some other vectors that are fed back. There are many ways in which you can use error back propagation neural network (EBPNN) for sequence-to-sequence tasks as the above diagram shows. Figure 2 shows hybrid deep learning, and we can feed in the input sequence to an BPNN and get a sequence of h vectors. We can then process these h vectors further.

3.4 *Data Acquisition*

Early-stage data processing strategy was data extraction using samples. The technical advances linked to these research instruments were used to allow them to be developed. These developments also made it easier to obtain visual readings of risk detection WBAN health monitoring data medical data. Samples have undergone a revolution as scientific science and conventional medicine are paired with two significant positions in the development of health care. By using hybrid deep learning, we have proposed the HDL model that can help us to generate the information about a risk patient and monitoring patient retention quality. Our basic goals are no more than to filter data, create improved efficient information and, also, develop a prediction model. In our first step, we are performing data preprocessing for proper extraction of datasets having some problem in them. In the second stage, model is particularly designed along with the use of data mining tools in it. Our data mining tools have capability to show the working performance. The proposed model that we have developed before and its specification have been determined, and, also, parameters are defined for further analysis. Now we have compared our model to other models, and we found that our proposed model is efficient in its own way. Now, here we are trying to complete with our final steps. The final stage comes that is evaluation. And in this final stage, we select the best parameters for right evaluation. Through this, we have obtained a result and output produced by the model. By this, we have critically increased the retention rate, and we have patient monitoring for behavior prediction:.

Our proposed model is to perform a number of steps through deep learning.

3.4.1 Collection of data

In this method for the collection of data, a number of data mining techniques have been used. By using the scraping tools to collect the available online information of WBAN data, we have prepared the data repository for suitable analysis of data. The data is collected from various

3.4.2 Preprocessing of data

In this phase of data preprocessing, we increase the efficiency of data prediction by using available techniques. These techniques will help to improve the quality of data preprocessing. The data preprocessing is indeed imperative to allow data available to user in a better way. This process converts the unwanted data to new modified datasets. If the data is not converted fully, then it reports the errors in the first step or even the results are not obtained accordingly, and then the whole step is performed again. In some cases, it might be that algorithms will be working in a proper way but the result will not be accordingly.

A number of steps to perform preprocessing dataset are as follows:

1. Here with this approach, the redundant data is removed and data is stored in a proper mode.
2. To format datasets, its arrangements are in a proper way which is highly required.
3. If the information is not according to the user knowledge, then the whole process will be repeated.
4. Now, it is needed to remove the unwanted information.
5. Here we can identify the missing values and note them to remember them for further usage.
6. There are some transactions performed by the user that are ambiguous in nature, so we will try to determine them.

4 Results Analysis

In this research work, we used MapReduce function based on HDL for complete separation of preprocessed exploration. The algorithm aimed at separating preprocessed MapReduce model for data processing and classification. One of the web application data collected completes the online streaming tool called flume which is founded on HADOOP ecosystem. MongoDB is used for saving the running status, and HDFS is used for storing the data. Split preprocessing information extraction performs exploration. This analyzes the framework in every verdict to improve the accuracy and focuses them on pattern equivalent with the negative context thesaurus and the positive situation thesaurus. Our analysis by using MapReduce data dictionary is composed of polarity and score. Lastly, we used Python language for envisioning the

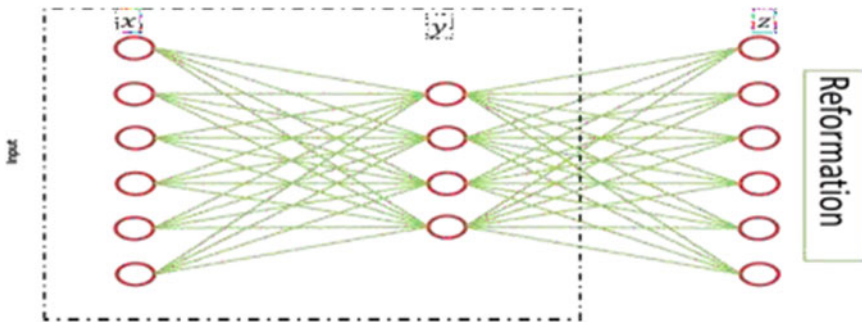


Fig. 1 BPNN for WBAN

total and polarization of the data. This technique is composed of an HDFS system created on the MapReduce model which classifies the ineffective anomaly detection in WBAN health monitoring data and score using hybrid learning-based MapReduce model. For analysis of live streaming, the WBAN health monitoring data was gathered using flume. We used Weka tools and HADOOP ecosystem for news data visualization, for news event personalization and for saving the status of information using the MongoDB.

It is a distributed, consistent and accessible service for well-organized gathering, combining, and for changing the position of huge volumes of WBAN data. It has a simple and flexible architecture created on delivering data flow. It is tough and fault tolerant with tunable dependability mechanism and numerous failover and rescue technique. It uses a simple and extensible data model that permits an online investigative application.

The person having heart rate more than 140 is more prone to heart attack disease. In Fig. 1, the Y-axis contains the number of person and X-axis contains the heart rate. The label-0 containing maximum number of persons is 6 whose heart rate is in the range of 130-135. Similarly, for label, the maximum number of persons is 8 whose heart rate is in the range of 160-170 which are more prone to heart attack disease. The precision score is given by formula as $(\text{True Positive}) / (\text{True Positive} + \text{False Positive})$, and recall is given as $(\text{True Positive}) / (\text{True Positive} + \text{False Negative})$. Figure 4 contains the graph of precision and recall which provides the average precision score of 0.9211, and the other graph is plotted for true positive rate and false positive rate which results that the area under the curve (AUC) score is 0.8921. AUC offers a cumulative quantity of performance across all probable classification thresholds.

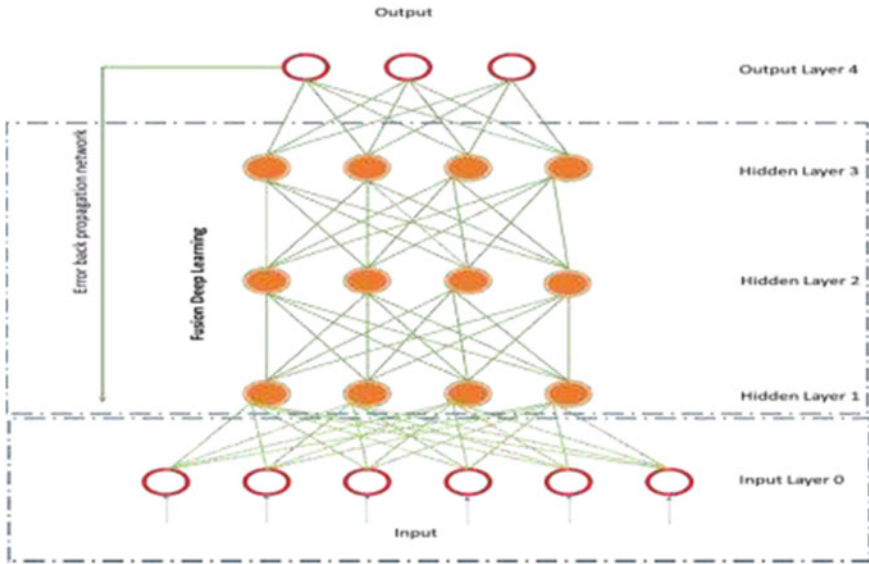


Fig. 2 Hybrid deep learning

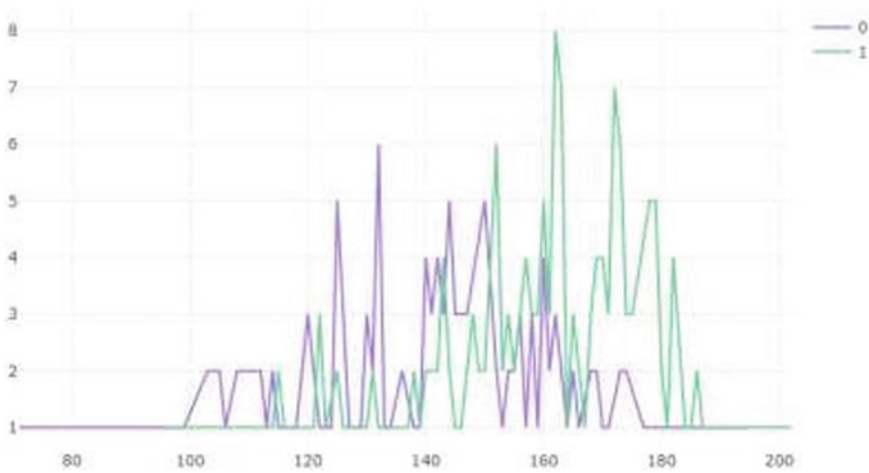


Fig. 3 Maximum heart rate achieved by labeled class 0 and 1

5 Conclusions and Future work

As the market for top-of-the-line methodologies of acquisition and processing of large sensed data is growing, the use of sensor and actuator networks in WBAN is wide speeding. WBAN services include the incorporation of intelligent healthcare

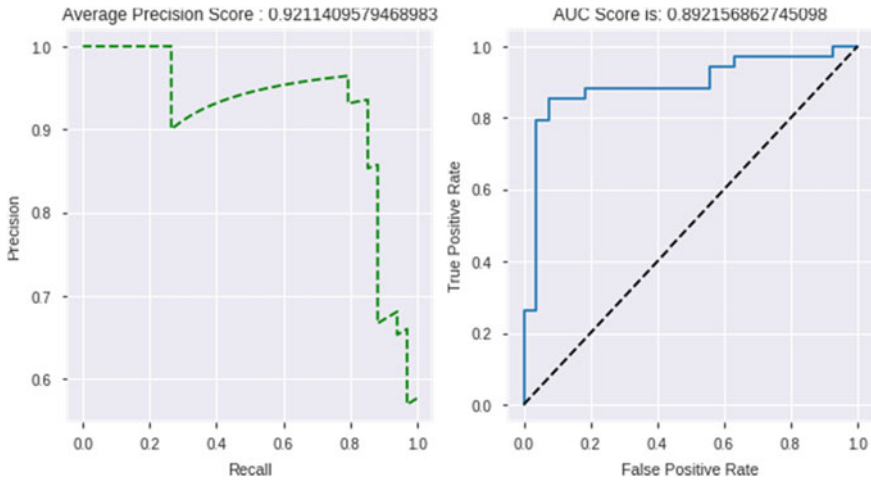


Fig. 4 Graph plotted for average precision score and AUC score

applications in everyday lives which extend the lives of society's citizens and enhance quality of life. The heterogeneous and different data types which are produced daily require efficient sensor acquisition and reliable, successful analysis in order to draw observations, forecasts and recommendations from sensory networks of sensors and actuators (e.g., wearables, carriers and other medical sensors). Deep learning is an effective tool used especially in these settings for the processing of large sensory data. We have also implemented many methods for the collection of data, such as wearables, samples and multiple sensors. In addition, the deep learning strategies studied have been related to known case uses for processing sensory data in medical practice. We assure that all relations are well-established listed sensory data collection strategies on the basis of current data generation technologies. We visited particular cases of usage and current deep learning tools for productive sensory data processing. We design model for risk detection in WBAN health monitoring data using hybrid deep learning. In future, we try to implement real-time environment for risk detection in wireless body sensor networks for health monitoring using hybrid deep learning.

References

1. L. He, C. Cao, Automated depression analysis using convolutional neural networks from speech. *Journal of Biomedical Informatics* **83**, 103–111 (2018)
2. Yang, L., Jiang, D., Xia, X., Pei, E., Oveneke, M.C. and Sahli, H., (2017). Multimodal Measurement of Depression Using Deep Learning Models. In *Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge (AVEC '17)*. Association for Computing Machinery, New York, NY, USA, 53–59. DOI:<https://doi.org/10.1145/3133944.3133948>

3. Cen, L., Yu, Z. L., Kluge, T., & Ser, W. (2018). Automatic System for Obstructive Sleep Apnea Events Detection Using Convolutional Neural Network. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). <https://doi.org/10.1109/embc.2018.8513363>
4. N. Yamanakkanavar, J.Y. Choi, B. Lee, MRI Segmentation and Classification of Human Brain Using Deep Learning for Diagnosis of Alzheimer's Disease: A Survey. *Sensors* **2020**(20), 3243 (2020)
5. Jo, T., Nho K., Saykin, A.J., (2019). Deep Learning in Alzheimer's Disease: Diagnostic Classification and Prognostic Prediction Using Neuroimaging Data. *Front Aging Neurosci.* 2019 Aug 20;11:220. <https://doi.org/10.3389/fnagi.2019.00220>. PMID: 31481890; PMCID: PMC6710444.
6. Siar, M., & Teshnehlab, M. (2019). Brain Tumor Detection Using Deep Neural Network and Machine Learning Algorithm. 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE). <https://doi.org/10.1109/iccke48569.2019.8964846>
7. Ahmed, O., & Brifceni, A. (2019). Gene Expression Classification Based on Deep Learning. 2019 4th Scientific International Conference Najaf (SICN). <https://doi.org/10.1109/sicn47020.2019.9019357>
8. Nazari, I., Tayara, H., & Chong, K. T. (2018). Branch Point Selection in RNA Splicing Using Deep Learning. *IEEE Access*, 1–1. doi:10.1109/access.2018.2886569
9. Tazawa, Y., Liang, K.C., Yoshimura, M., Kitazawa, M., Kaise, Y., Takamiya, A., Kishi, A., Horigome, T., Mitsukura, Y., Mimura, M., Kishimoto, T., (2020). Evaluating depression with multimodal wristband-type wearable device: screening and assessing patient severity utilizing machine-learning. *Heliyon*. 6. e03274. 10.1016/j.heliyon.2020.e03274
10. McGinnis, R. S., McGinnis, E. W., Hruschak, J., Lopez-Duran, N. L., Fitzgerald, K., Rosenblum, K. L., & Muzik, M. (2018). Wearable sensors and machine learning diagnose anxiety and depression in young children. 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI). <https://doi.org/10.1109/bhi.2018.8333455>
11. G.B. Papini, P. Fonseca, M.M. van Gilst et al., Wearable monitoring of sleep-disordered breathing: estimation of the apnea-hypopnea index using wrist-worn reflective photoplethysmography. *Sci Rep* **10**, 13512 (2020). <https://doi.org/10.1038/s41598-020-69935-7>
12. G. Surrel, A. Aminifar, F. Rincón, S. Murali, D. Atienza, Online Obstructive Sleep Apnea Detection on Medical Wearable Sensors. *IEEE Transactions on Biomedical Circuits and Systems* **12**(4), 762–773 (2018). <https://doi.org/10.1109/TBCAS.2018.2824659>
13. Zhou, P., Huang, L., Zhao, Q., Xiao, W., Li, S., (2019). A Domestic Diagnosis System for Early Restless Legs Syndrome Based on Deep Learning. *Zhongguo Yi Liao Qi Xie Za Zhi*. 2019 Mar 30;43(2):79-82. Chinese. <https://doi.org/10.3969/j.issn.1671-7104.2019.02.001>. PMID: 30977599.
14. Rajawat A.S., Upadhyay P., Upadhyay A. (2021) Novel Deep Learning Model for Uncertainty Prediction in Mobile Computing. In: Arai K., Kapoor S., Bhatia R. (eds) *Intelligent Systems and Applications*. *IntelliSys 2020*. *Advances in Intelligent Systems and Computing*, vol 1250. Springer, Cham. https://doi.org/10.1007/978-3-030-55180-3_49
15. K. Barhanpurkar, A. S. Rajawat, P. Bedi and O. Mohammed, "Detection of Sleep Apnea & Cancer Mutual Symptoms Using Deep Learning Techniques," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2020, pp. 821-828, <https://doi.org/10.1109/i-smac49090.2020.9243488>.
16. A. Singh Rajawat and S. Jain, "Fusion Deep Learning Based on Back Propagation Neural Network for Personalization," 2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India, 2020, pp. 1-7, <https://doi.org/10.1109/idea49133.2020.9170693>.
17. A. S. Rajawat, O. Mohammed and P. Bedi, "FDLM: Fusion Deep Learning Model for Classifying Obstructive Sleep Apnea and Type 2 Diabetes," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2020, pp. 835- 839, <https://doi.org/10.1109/i-smac49090.2020.9243553>.

18. A. S. Rajawat and A. R. Upadhyay, "Web Personalization Model Using Modified S3VM Algorithm For developing Recommendation Process," 2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India, 2020, pp. 1-6, <https://doi.org/10.1109/idea49133.2020.9170701>.
19. Sampurna Mandal;Swagatam Biswas;Valentina E. Balas;Rabindra Nath Shaw; Ankush Ghosh, "Motion Prediction for Autonomous Vehicles from Lyft Dataset using Deep Learning" 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA) 30-31 Oct. 2020, Page(s):768 – 773, <https://doi.org/10.1109/iccca49541.2020.9250790>
20. Sampurna Mandal;Valentina E. Balas;Rabindra Nath Shaw; Ankush Ghosh, "Prediction Analysis of Idiopathic Pulmonary Fibrosis Progression from OSIC Dataset" 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON), 2-4 Oct. 2020,Page(s):861 – 865, <https://doi.org/10.1109/gucon48875.2020.9231239>
21. Kumar M., Shenbagaraman V.M., Shaw R.N., Ghosh A. (2021) Predictive Data Analysis for Energy Management of a Smart Factory Leading to Sustainability. In: Favorskaya M., Mekhilef S., Pandey R., Singh N. (eds) Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering, vol 661. Springer, Singapore. https://doi.org/10.1007/978-981-15-4692-1_58
22. AbdallaAlameen, Ashu Gupta, Optimization Driven Deep Learning Approach for Health Monitoring and Risk Assessment in Wireless Body Sensor Networks
23. R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, R.X. Gao, Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing* **115**, 213–237 (2019). <https://doi.org/10.1016/j.ymssp.2018.05.050>
24. Xia, M., Zheng, X., Imran, M., & Shoaib, M. (2020). Data-driven prognosis method using hybrid deep recurrent neural network. *Applied Soft Computing*, 106351. doi:10.1016/j.asoc.2020.106351
25. L. Ding, W. Fang, H. Luo, P.E.D. Love, B. Zhong, X. Ouyang, A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory. *Automation in Construction* **86**, 118–124 (2018). <https://doi.org/10.1016/j.autcon.2017.11.002>
26. Jones, V., Gay, V., & Leijdekkers, P. (2010). Body Sensor Networks for Mobile Health Monitoring: Experience in Europe and Australia. 2010 Fourth International Conference on Digital Society. <https://doi.org/10.1109/icds.2010.41>
27. A. Alameen, A. Gupta, Optimization Driven Deep Learning Approach for Health Monitoring and Risk Assessment in Wireless Body Sensor Networks. *International Journal of Business Data Communications and Networking (IJBDCN)* **16**(1), 70–93 (2020). <https://doi.org/10.4018/IJBDCN.2020010105>
28. X. Liu, M. Zhao, A. Liu, K.K.L. Wong, Adjusting forwarder nodes and duty cycle using packet aggregation routing for body sensor networks. *Information Fusion* **53**, 183–195 (2020). <https://doi.org/10.1016/j.inffus.2019.06.020>
29. Chen, H., Wang, J., Dong, X., & Zhao, C. (2020). Security design of ECG telemonitoring systems. 2020 International Conference on Computer Engineering and Application (ICCEA). <https://doi.org/10.1109/iccea50009.2020.00154>
30. S.J. Hussain, M. Irfan, N.Z. Jhanjhi, K. Hussain, M. Humayun, Performance Enhancement in Wireless Body Area Networks with Secure Communication. *Wireless Personal Communications* (2020). <https://doi.org/10.1007/s11277-020-07702-7>
31. Noseworthy, P. A., Attia, Z. I., Brewer, L. C., Hayes, S. N., Yao, X., Kapa, S., ... Lopez-Jimenez, F. (2020). Assessing and Mitigating Bias in Medical Artificial Intelligence: The Effects of Race and Ethnicity on a Deep Learning Model for ECG Analysis. *Circulation: Arrhythmia and Electrophysiology*. <https://doi.org/10.1161/circep.119.007988>
32. Shankar, K., Zhang, Y., Liu, Y., Wu, L., & Chen, C.-H. (2020). Hyperparameter Tuning Deep Learning for Diabetic Retinopathy Fundus Image Classification. *IEEE Access*, 1–1. doi:10.1109/access.2020.3005152
33. Jiang, H., Yang, K., Gao, M., Zhang, D., Ma, H., & Qian, W. (2019). An Interpretable Ensemble Deep Learning Model for Diabetic Retinopathy Disease Classification. 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). <https://doi.org/10.1109/embc.2019.8857160>

34. X. Li, L. Shen, M. Shen, F. Tan, C.S. Qiu, Deep learning based early stage diabetic retinopathy detection using optical coherence tomography. *Neurocomputing* (2019). <https://doi.org/10.1016/j.neucom.2019.08.079>
35. F. Ali, S. El-Sappagh, S.M.R. Islam, D. Kwak, A. Ali, M. Imran, K.-S. Kwak, A Smart Healthcare Monitoring System for Heart Disease Prediction Based On Ensemble Deep Learning and Feature Fusion. *Information Fusion* (2020). <https://doi.org/10.1016/j.inffus.2020.06.008>
36. Rathore, H., Al-Ali, A., Mohamed, A., Du, X., & Guizani, M. (2019). A Novel Deep Learning Strategy for Classifying Different Attack Patterns for Deep Brain Implants. *IEEE Access*, 1–1. doi:10.1109/access.2019.2899558
37. S. Dash, S.K. Shakyawar, M. Sharma, Big data in healthcare: management, analysis and future prospects. *J Big Data* **6**, 54 (2019). <https://doi.org/10.1186/s40537-019-0217-0>

AI and IOT-Based Model for Photovoltaic Power Generation



Harshit Kumar Huneria, Pavan Yadav, Rabindra Nath Shaw, D. Saravanan,
and Ankush Ghosh

Abstract In recent years, it is seen that the amount of energy consumption is increasing rapidly as everyone is looking for development and economical growth; also, we have limited resources fossil fuel, electrical energy, and mechanical energy which are still maximum source of energy of the world. That is why electrical energy is one of the major components for fulfilling the needs of the world. Many analysts tried to create electrical energy in the ground of renewable sources using different sensors, so that they can reduce the use of fossil fuel and also for the fulfillment of electrical energy for different appliances. In this chapter, we are focusing on the same challenges and making it work to generate an electrical energy using sensors, i.e., photovoltaic plate, which is based on Internet of things (IoT) system for different appliance. Energy developed by photovoltaic plate is connected to the power storage circuit. Here, total power generated to exploit to different AI models approval is done through the measurable boundaries like RMSE and RRMSE.

Keywords Artificial intelligence · Internet of things · Photovoltaic plates · RMSE

1 Introduction

1.1 Significance of Renewable Energy

In present scenario, inexhaustible aquifers of energy assume a crucial function to fulfill the interest of power energy age. Request of power is the critical element to regard for human development in order of economical and financial development. Expanding request of power and divulgence of smart grid system is pulled in home

H. K. Huneria · P. Yadav · R. N. Shaw · D. Saravanan
Department of Electrical, Electronics and Communication Engineering, Galgotias University,
Greater Noida, India
e-mail: r.n.s@ieee.org

A. Ghosh (✉)
School of Engineering and Applied Sciences, The Neotia University, Sarisha, West Bengal, India

energy the board framework (HEMS) to lessen the utilization of power. Electric force plants are going through significant change age move from enormous unified electric force plants to little sustainable decentralized force stations.

Renewable energy is a sensible power source gathered from inexhaustible capital, and they are normal viewed on earthborn timelines. Non-renewable orgone generated from link that were not refresh in duration of life maximum energy more that 50% is progressed by non-renewable energy. Hence, it was adjusted that as soon as we might be facing the shortage of petroleum energy. The invention of electricity from the renewable source is wind, and sun energy is high end. In first half of 2020, the amount of energy produces from wind and solar which was 10% of the total energy production. For requirement of electrical energy in build up their lack of hygienic energy foundation, each and every possible energy (small or large) fulfills the need of electrical energy. The store no renewable source is build up in the world. Hence, many scholars tried to give and generate electricity from renewable energy by sensors in right to over lift store electrical energy for essential things [1, 2].

1.2 Importance of Machine Learning

In present days, machine learning language plays an important role in decipher various issues and troubles in many field in real world. Some sort of machine learning techniques that are useful to deliver valuable information related to sustainable source of energy for generation as well as integration are artificial neural network (ANN), linear regression, polynomial regression are some of them [3–6]. During the past few decades, application of artificial intelligence strategies has quickly extended in various renewable power generation cases similar to less upfront capital cost, constrained storage capacity, and geological confinement.

1.3 Pros of Machine Learning

- Self-learning machine can be delegated using machine learning to perform complex tasks [7].
- No human intervention needed.
- Multi-dimensional and multi-variety data any easily handled.
- Easily identify patterns and trends.

1.4 Cons of Machine Learning

- High error susceptibility
- Data acquisition

- Misinterpretation of results

2 Literature Survey

2.1 Frame Work

Present day, machine learning technology is playing an essential role in the overall development, as current it used in every sector of field to reduce dependence on man power [8–10].

The purpose of this execution is to exhibit the smart and brilliant of the machine learning technique. The model on which we are working is for power generation from a renewable energy resource [11–14]. Here, the resource is sunlight as we are using photovoltaic panel for power generation. This system is connected to Arduino which is essential and makes our task to connect and deliver output data of the model to the server using Wi-Fi module. The collective data is later stored for the validation of the system.

2.2 Validation

For validation of the generated power output, couple of methods are used. Root mean square error (RMSE) and relative mean square error (RRMSE) are used to distinguish the efficient model [15–18]. The accompanying conditions are utilized to measure execution measurement boundaries which are as follows:

$$\text{rmse} = \sqrt{\text{mean}((z_{\text{pre}} - z)^2)} \quad (1)$$

$$\text{rrmse} = \text{rmse}/\text{mean}(z) \quad (2)$$

where z_{pre} is the predicted value and z is the actual value.

3 Methodology

3.1 Computing and Processing the Data

For that we create a circuit on MATLAB/Simulink, where we connected a solar panel block with an input intake of constant temperature of 25 °C and a constant irradiance of 1000 W/m². This solar panel block is further connected with a bypass diode, current

measuring, and voltage measuring blocks [19–21]. The output of current measuring and voltage measuring block goes to the scope, respectively, and its outlet value under goes to a product block as $P = V * I$ which is connected to power measuring instrument [22–24]. The data for the simulation is export in the form of.csv file to workspace. Solar panel generally produces 12 V and differs as indicated by the size of panel.

3.2 Export File to Server

ESP8266 is utilized for the connection in between information and server. AT command set is used to make it work according. Furthermore, it is outfitted with installed antenna just as help TCP/IP Convention for good connectivity.

3.3 Modeling

Linear regression and polynomial regression of ANN are utilized for the exactness of the model, just to figure out which model is appropriate for the power yield [25–27]. Linear regression and polynomial regression in ANN are utilized in approval of model and factual strategies, where RMSE and RRMSE are utilized for validation (Fig. 1).

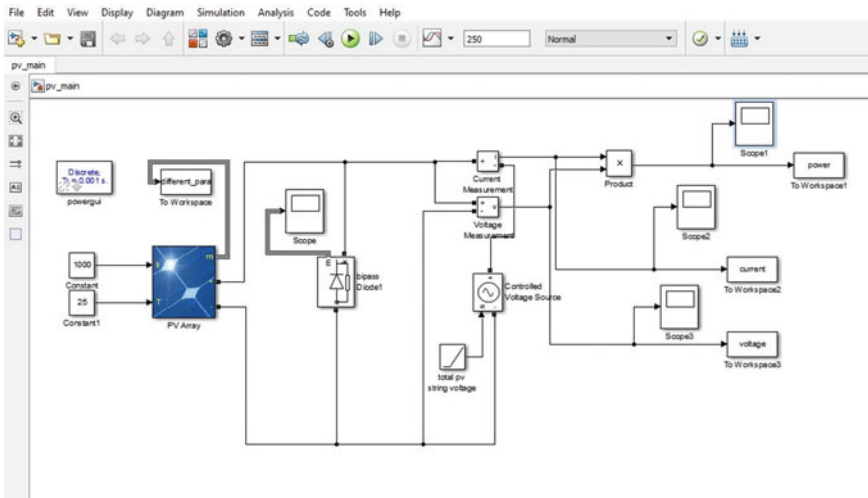


Fig. 1 Model for data processing

Fig. 2 Sample data used for simulation

time	Iout	Vout	Power
0	6.20614	0	0
0.001	6.204834	0.001	0.006205
0.002	6.20445	0.002	0.012409
0.003	6.204337	0.003	0.018613
0.004	6.204302	0.004	0.024817
0.005	6.204291	0.005	0.031021
0.006	6.204286	0.006	0.037226
0.007	6.204284	0.007	0.04343
0.008	6.204282	0.008	0.049634
0.009	6.20428	0.009	0.055839
0.01	6.204279	0.01	0.062043
0.011	6.204277	0.011	0.068247
0.012	6.204276	0.012	0.074451
0.013	6.204274	0.013	0.080656
0.014	6.204272	0.014	0.08686
0.015	6.204271	0.015	0.093064

3.4 Software Used

- **MATLAB:**
We used this software to construct our circuit on MATLAB/Simulink and collect the informative data from the Simulink and export it to editor. In editor, we used the import data from the Simulink stored by the file name.csv. and that file data is further plotted in the form of graphs. Data is also used in the validation of working model.
- **Arduino IDE:**
This software is quite familiar to us, IDE is based on C and C++ programming language and most of the required components were available. It is utilized to compose and compile the code for Arduino so that we can connect are model together any transfer the informative data to the server using Wi-Fi module.
- **Proteus 8 simulator:**
Proteus is extraordinary programming and simulation program for different circuit design. It is utilized to perform simulation (Figs. 2, 3, 4, 5, 6 and 7).

From the plot that the polynomial value RMSE is less than the linear value of RMSE. Therefore, polynomial regression model is showing more accurate than linear regression model [28].

4 Conclusions

The model that has been implemented for power generation based on IOT can supply energy to different appliances. Here, different methods are used to simulate the model

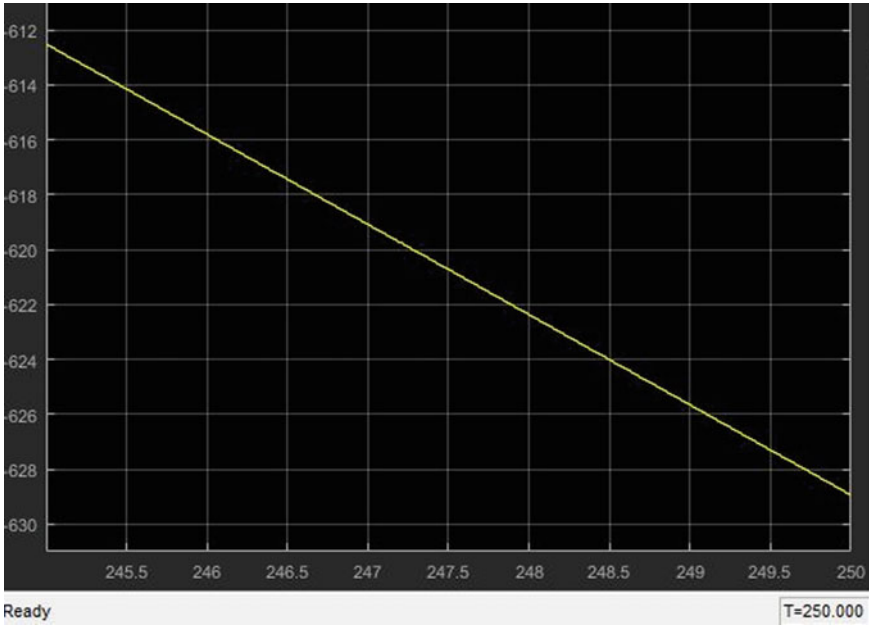


Fig. 3 Current output generated by solar cell with deferent time interval

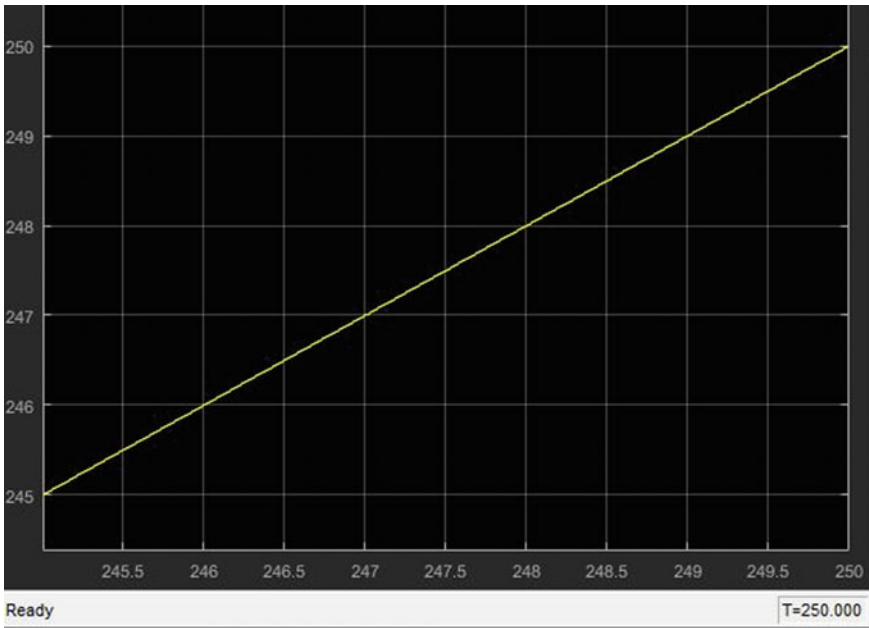


Fig. 4 Voltage output generated by solar cell with deferent time interval

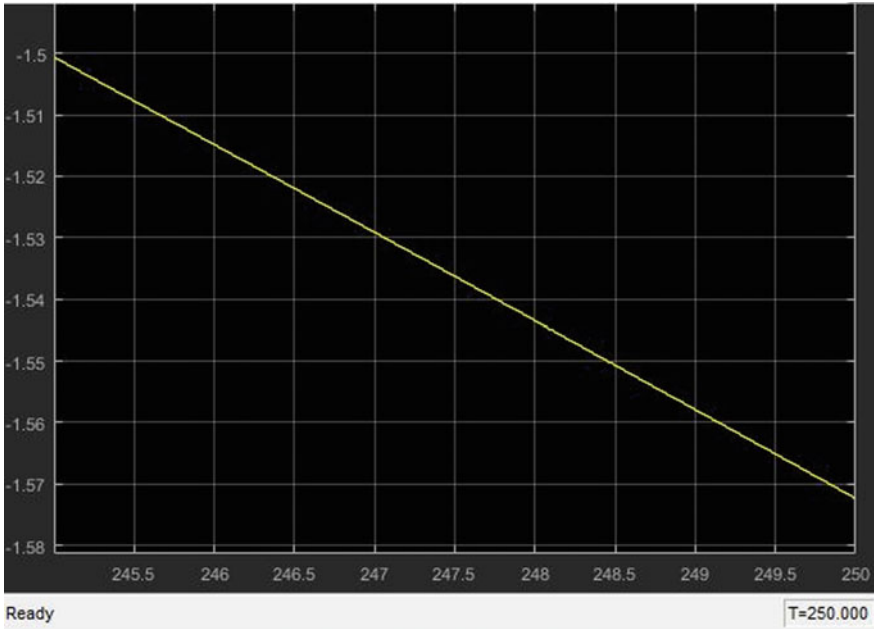


Fig. 5 Power output calculated with deferent time interval

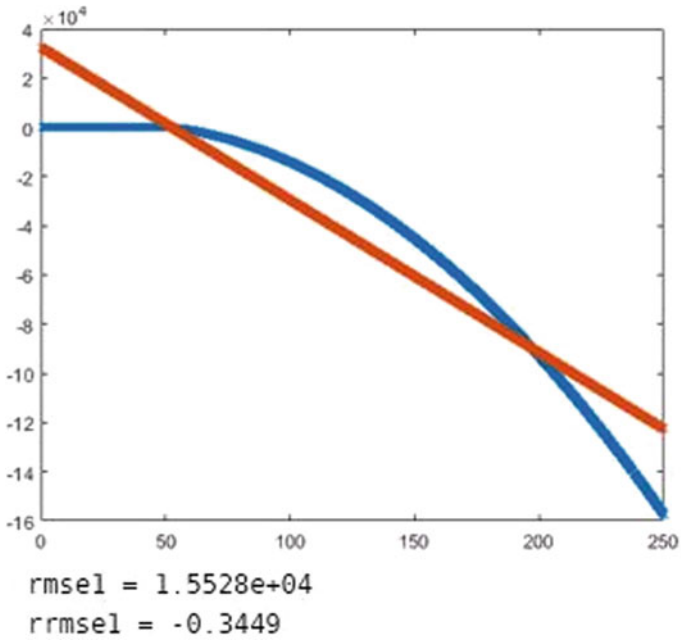


Fig. 6 Linear regression model on power output with actual value versus predicted value

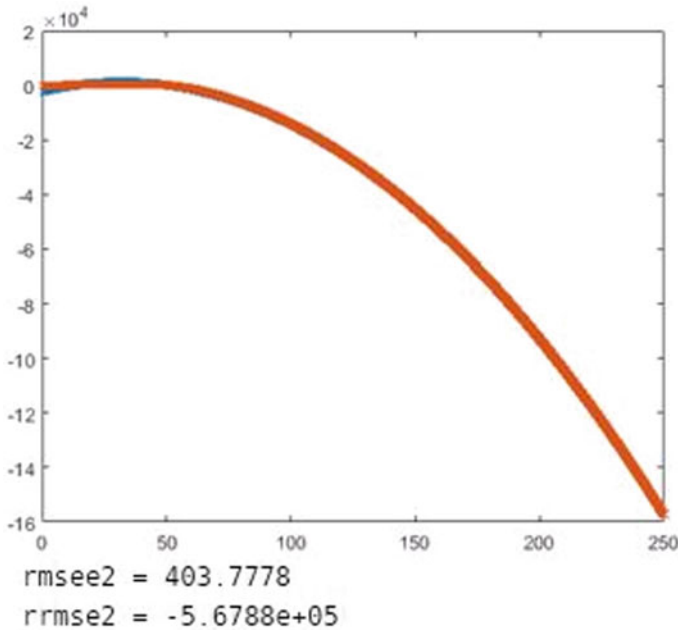


Fig. 7 Polynomial regression model on power output with actual value versus predicted value

to get a valid result. Simulation of photovoltaic panel is done through MATLAB, while IOT and Arduino-based simulation are done on PROTEUS 8. For the validation of the model certain method, RMSE and RRMSE have been performed, which shows that both polynomial regression of ANN and linear regression of ANN of given decent predicted value of z_{pre} . So far from collective data, we observed that polynomial regression is more effective method in predicting the value of the correspondent output.

References


1. V. Puri, S. Jha, R. Kumar, I. Priyandarshini, et al. A hybrid artificial intelligence and internet of things model for generation of renewable resource of energy. *IEEE Access* (2019)
2. P. Garcia, C.A. Garcia, L.M. Fernandez, F. Llorens, F. Jurado, ANFIS-based control of a grid connected hybrid system integrating renewable energies, hydrogen and batteries. *IEEE Trans. Ind. Inf.* (2014)
3. Y. Belkhier, A. Achour, R.N. Shaw, Fuzzy passivity-based voltage controller strategy of grid-connected PMSG-based wind renewable energy system, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India (2020), pp. 210–214. <https://doi.org/10.1109/ICCCA49541.2020.9250838>
4. G. Kapoor, R.N. Shaw, Identification of faults in parallel transmission lines using discrete Meyer wavelet, in *2020 IEEE 5th International Conference on Computing Communication*

- and Automation (ICCCA)*, Greater Noida, India (2020), pp. 736–741. <https://doi.org/10.1109/ICCCA49541.2020.9250831>
5. G. Kapoor, R.N. Shaw, A MMG indices-based scheme for fault detection in series-compensated power transmission system, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India (2020), pp. 742–747. <https://doi.org/10.1109/ICCCA49541.2020.9250714>
 6. I. Das, R.N. Shaw, S. Das, Performance analysis of wireless sensor networks in presence of faulty nodes, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India (2020), pp. 748–751. <https://doi.org/10.1109/ICCA49541.2020.9250724>
 7. S. Mandal, S. Biswas, V.E. Balas, R.N. Shaw, A. Ghosh, Motion prediction for autonomous vehicles from Lyft dataset using deep learning, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India (2020), pp. 768–773. <https://doi.org/10.1109/ICCCA49541.2020.9250790>
 8. B. Kantha, R.N. Shaw, A. Ghosh, Influence of annealing temperatures in improving the sensitivity of Pd modified WO₃ nanoparticles based hydrogen sensor, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India (2020), pp. 780–784. <https://doi.org/10.1109/ICCCA49541.2020.9250788>
 9. R. Tripathi, S. Agrawal, R.N. Shaw, Impact of number of collector on energy profile from concentrated fully covered hybrid photovoltaic (CHPV) system, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India (2020), pp. 802–806. <https://doi.org/10.1109/ICCCA49541.2020.9250896>
 10. D. Saravanan, A. Hasan, A. Singh, H. Mansoor, R.N. Shaw, Fault prediction of transformer using machine learning and DGA, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India (2020), pp. 1–5. <https://doi.org/10.1109/GUCON48875.2020.9231086>
 11. G. Kapoor, R.N. Shaw, Protection of dual-circuit transmission line integrated with wind-park using fast Walsh-Hadamard transform, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India (2020), pp. 480–485. <https://doi.org/10.1109/GUCON48875.2020.9231092>
 12. I. Das, R.N. Shaw, S. Das, Analysis of effect of fading models in wireless sensor networks, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India (2020), pp. 858–860. <https://doi.org/10.1109/GUCON48875.2020.9231201>
 13. G. Kapoor, A. Mishra, R.N. Shaw, S. Ajmera, Complex faults categorization technique for series-compensated transmission lines connected with wind-turbines based on DFVHT, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India (2020), pp. 703–708. <https://doi.org/10.1109/GUCON48875.2020.9231152>
 14. B. Kantha, R.N. Shaw, A. Ghosh, The effect of catalytic metal contact and surface modification on hydrogen sensing performance of nano WO₃/SiO₂/Si sensor, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India (2020), pp. 866–869. <https://doi.org/10.1109/GUCON48875.2020.9231258>
 15. S. Mandal, V.E. Balas, R.N. Shaw, A. Ghosh, Prediction analysis of idiopathic pulmonary fibrosis progression from OSIC dataset, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India (2020), pp. 861–865. <https://doi.org/10.1109/GUCON48875.2020.9231239>
 16. R.N. Shaw, P. Walde, A. Ghosh, IOT based MPPT for performance improvement of solar PV arrays operating under partial shade dispersion, in *2020 IEEE 9th Power India International Conference (PIICON)*, SONEPAT, India (2020), pp. 1–4. <https://doi.org/10.1109/PIICON49524.2020.9112952>
 17. G. Kapoor, R.N. Shaw, SWT-based fault recognition technique for six-phase transmission system, in *2020 IEEE International Conference on Computing, Power and Communication*

- Technologies (GUCON)*, Greater Noida, India (2020), pp. 328–333. <https://doi.org/10.1109/GUCON48875.2020.9231216>
18. S. Paul, J.K. Verma, A. Datta, R.N. Shaw, A. Saikia, Deep learning and its importance for early signature of neuronal disorders, in *2018 4th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India (2018), pp. 1–5. <https://doi.org/10.1109/CCAA.2018.8777527>
 19. D. Saravanan, D. Singh, S. Sinha, R.N. Shaw, K. Prabu, A comprehensive study of well-known maximum power point tracking techniques, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India (2020), pp. 829–837. <https://doi.org/10.1109/GUCON48875.2020.9231219>
 20. R.S. Somalwar, S.G. Kadwane, R.N. Shaw, Frequency estimation by recursive least square in active islanding method for microgrid, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India (2020), pp. 116–123. <https://doi.org/10.1109/GUCON48875.2020.9231206>
 21. M.D. Siddique, R.N. Shaw, et al. 2020 Switched-capacitor based seven-level triple voltage gain boost inverter (7L-TVG-BI), in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India (2020), pp. 848–852. <https://doi.org/10.1109/GUCON48875.2020.9231182>
 22. I. Das, R.N. Shaw, S. Das, Analysis of energy consumption in dynamic mobile ad hoc networks, in *Advances in Intelligent Systems and Computing*, vol. 1049 (Springer, Berlin, 2020), pp. 235–243. https://doi.org/10.1007/978-981-15-0132-6_15
 23. R.N. Shaw, P. Walde, A. Ghosh, A new model to enhance the power and performances of 4×4 PV arrays with puzzle shade dispersion. *Int. J. Innov. Technol. Exploring Eng.* **8**(12), 456–465 (2019). <https://doi.org/10.35940/ijitee.L3338.1081219>
 24. M. Saxena, R.N. Shaw, J.K. Verma, A novel hash-based mutual RFID tag authentication protocol, in *Advances in Intelligent Systems and Computing*, vol. 847. (Springer, Berlin, 2019), pp. 1–12. https://doi.org/10.1007/978-981-13-2254-9_1
 25. R.N. Shaw, D. Basu, P. Walde, A. Ghosh, Effects of solar irradiance on load sharing of integrated photovoltaic system with IEEE standard bus network. *Int. J. Eng. Adv. Technol.* **9**(1), 424–429 (2019). <https://doi.org/10.35940/ijeat.A3188.109119>
 26. I. Das, R.N. Shaw, S. Das, Analysis of energy consumption of energy models in wireless sensor networks, in *Innovations in Electrical and Electronic Engineering*, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh. *Lecture Notes in Electrical Engineering*, vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_57
 27. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering*, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh. *Lecture Notes in Electrical Engineering*, vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
 28. I. Das, R.N. Shaw, S. Das, Location-based and multipath routing performance analysis for energy consumption in wireless sensor networks, in *Innovations in Electrical and Electronic Engineering*, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh. *Lecture Notes in Electrical Engineering*, vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_59

Combating Challenges in the Construction Industry with Blockchain Technology



Priyanka Singh , Debarati Sammanit, Paritosh Krishnan,
Krishna Mohan Agarwal, Rabindra Nath Shaw, and Ankush Ghosh

Abstract Blockchain technology serves a breakthrough for many solutions like that of poor payment facilities, lesser productivity, trust among different levels of the industry in sharing of the gathered field reports, or the commands of the higher authority in the construction industry. Thus, the decentralized ledger system has facilitated an automatic chain of the flow of data or information among the peers and the subordinates. The aim is to make a digital world more trustable giving a new open arena to a better security solution, resilient nature and the enhancement of the systems. It is already known that industry progress upon its embracing of newer technologies that bind a better relationship with the customers and the giving of the best quality project basing on the low cost of trade estimates; the blockchain facilitates the use of smart contracts, agreements with the party directly without the involvement of a third party. The new version of the industries is supposedly aiming towards more digitized industrialization with more involvement in construction management with building information modelling (BIM). The aim is to provide a pre-step towards the use cases of blockchain technology in the construction industry after the reviving of the industry using the utmost digitization.

P. Singh · D. Sammanit

Department of Civil Engineering, Amity School of Engineering & Technology, Amity University
Uttar Pradesh, Noida, India

P. Krishnan

Department of Information Technology, Amity School of Engineering & Technology, Amity
University Uttar Pradesh, Noida, India

K. M. Agarwal

Department of Mechanical Engineering, Amity School of Engineering & Technology, Amity
University Uttar Pradesh, Noida, India

R. N. Shaw

Department of Electrical, Electronics and Communication Engineering, Galgotias University,
Noida, India

A. Ghosh (✉)

School of Engineering and Applied Sciences, The Neotia University, Sarisha, West Bengal, India

Keywords Blockchain technology · Construction industry · Infrastructure · Smart contracts · BIM

1 Introduction

The cryptocurrency in the present time has gained quite a popularity and is also getting trusted; an instance of the cryptocurrency is the bitcoin that had its unknown emergence in the year 2008 that is based on the blockchain methodology; the distributed ledger system that is involved in the blockchain helps in the segregation of the type of work involved that starts from the designing of a plan, its approval, the managing of the contracts between the stakeholders, the working of the plan, the plan execution before it gives the master project market-ready [1–3]. As these step-by-step processes go on with every new transaction being made a new block is added to the system of the distributed networks that get sanctioned and validated by the other blocks; thus, the need for a third party for the validation is avoided. The features that could be listed in as the fast-changing type of the industry need anonymity, the data transparency, the auditable nature involved in the smooth enabling of the blockchains in the blockchain network, or the nodal networks [4–6].

Rather than the outsider mediation, the exchanges conveyed by the excavators or the genuine holders of these squares are known to have constructed trust among the clients that their information is protected and not altered effectively; presently, one may impel on the reality of the nodal disappointment the response to it is the nodal frameworks that have the idea of the Internet of things (IoT) is each of the appropriated frameworks that help in restoring the information to the squares. The keen agreements that get conveyed in the blockchain that incorporates the proprietors and advertisers and the contractual workers bound to and by the conditions of the agreements that are approved by the gatherings for it getting remembered for the chain [7–9].

2 Challenges in the Construction Industry

For the undertaking execution in the development business, it ranges from the arrangement endorsement to the common assent of gatherings, temporary workers, the providers alongside the outside and inward group associated with the venture the board generally speaking [10, 11]. The condition of achievement must be acclaimed in a venture when the correct exchange that might be utilizing the record frameworks that prompt a definitive structure of the trust among the gatherings. These gatherings meet up combined under the shelter of the blockchain framework. The blockchain prompts the keeping up of the Ledger-based exchange having specific IDs that thus helps in keeping up the record of bills, the execution of the arrangement, and can be investigated after the culmination of the venture. The mechanisms like the smart

contracts, Ethereum-based contracts all helping to make up a legal foundation that in turn assures the quality planning, estimation of the cost, and giving a deadline for the project completion [3, 12].

The different challenges thus if with the implication of the blockchain-based systems will insight a step towards the sorting of the most controversial industry of mankind.

3 Blockchain Architecture

The distributed ledger system is a sequence of blocks with not a single source of data flow but multiple entry points of the data at different points or the nodes that hold good the transaction details the contract-based executions in the form of blocks, the mother block, or the first block to the chain is the genesis block and next blocks continue to add on with the updation [13]. The block is comprised of the block version, its block body. The Merkle tree root hash with its hashing features adds the hash values to all transactions, the timestamp sets the timing; the date and time, the bits giving to a threshold target of a valid hashing and importantly the 4 bytes nonce that keeps increasing with every new addition to the block. The basis of which is an asymptomatic cryptographic style followed using the digital signatures [14, 15].

4 Blockchain and Data Privacy

The blockchain gives a thorough element of permanence that thusly ensures the security of the information, the chain shaped with the purposes of passage for the stream data sends among the duplicates of each information to each obstruct, that when disagrees with any of a square during the exchange approval will be gotten fake and can't be added to the organization. However, the security issue can't be precluded as the decentralized record framework encourages the sharing of a duplicate of the apparent multitude of factors to the clients obscure in this way the darknet action may sneak in as of the deception to mine through the qualities, yet that stands essentially difficult to present any misbehaviour in the blockchain the excavators who mine the following square in the chain need the claiming of more than half of the blockchain that sabotages the strategy of blockchains to be a disseminated record [16, 17].

5 Features of Blockchain

Out of the blue, trade in the blockchain framework can be coordinated between any two companions (P2P) without approval by the central office. Consequently, blockchain would altogether be able to diminish the worker costs (considering the

consequences, additionally, the action cost) and moderate the show bottlenecks at the central worker.

Decentralization. In standard united trade structures, each trade ought to be endorsed through the central trusted in the association (e.g. the public bank) coming about the cost and the introduction bottlenecks at the central workers.

Persistency. Since all the trades spreading over the framework ought to be attested and recorded in squares circled in the whole framework, it is practically hard to change. Additionally, each conveyed square would be endorsed by various centre points and trades would be checked. So, any contortion could be distinguished with no issue.

Anonymity. Each customer can connect with the blockchain associate with a delivered address. Further, a customer could make various conveys to keep up a key good way from character introduction. There isn't, now any central social occasion keeping customers' shrouded information. This framework spares a particular proportion of security on the trades associated with the blockchain.

Auditability. Since all the trades on the blockchain are affirmed and recorded with a timestamp, customers can without doubt affirm and finish the records getting to any centre point in the suitable arrange. In bitcoin blockchain, each trade could be followed to past trades iteratively. It improves the perceptibility and the straightforwardness of the data set aside in the blockchain.

6 Consensus Calculation

With the agreeing of the various partners included or the excavators in the chains, the agreement calculation is a trustable also in the ramifications of the circulated record-based frameworks. The hidden conventions help assemble an organization of hubs with no focal hub except for appropriated frameworks.

Utilizing systems of verification of work which reference can be involved from the bitcoin network that was made in 2008. The legitimate age of squares added to the organization with the hashing procedures [18, 19].

The agreement calculation has a more noteworthy method for a productive chain framework guaranteeing better security and accommodation.

7 Consensus Methodology

The regular gathering of issues that look like that of the byzantine General issues has a basic idea of agreeing before arriving at a choice is like the agreement calculation. Agreeing to a commonality, i.e. agreement. A few stages are followed to agree likewise in the blockchain.

PoW—the Proof of Work dimes as a technique in the cryptographic money organization of bitcoins, the intention is to give a vehicle of arrangement dependent on

the recorded exhibit of the decentralized record of exchanges that are likewise put away in the framework which we call the blockchain. The permitting of any of the exchanges to be avoided by the BC network requires a ton of approval on some portion of each digger to approve for agreeing to prohibit the square there. With each purpose of the chain under the perception of the excavators computing the hash estimation of the squares in line alongside the header. Hence, during this cycle of mining the nonce, the focussed on the estimation of the square header is gotten by the significant figuring that are approved by different hubs by affirming the mined hashing values. On crucial occasions a few squares may produce comparative determined nonce esteems by then the standard of the longest chain is followed to let the progression of the exchange records to the organization [20–22].

Additionally, the proof of stake (PoS) serves as a royal option in contrast to the above procedure to arrive at the shared concurring of the legitimacy of exchanges, there are plentiful opportunities for diggers of all sort to be in an organization whose primary point is to control the header of the chain for affecting the progression of the information on his made line of the stream of the chain along these lines to keep away from this widespread peril the conceivable organization centres around permitting the excavators just to get to dependent on their capital or records; however, a one-sided framework comes up yet the best way to limit the danger of fabrication the means have been taken up by various digital money utilizing motors like the Peercoin. The more coins partner mines and expanding his benefits online gives him a more prominent likelihood to impact the mining of the following greater square and adding to the organization alongside peer agreement of the distinctive nonce [23, 24].

8 Decentralized Record in the Construction Business

The most isolated industry ever that reaches to bring various segments like the development of the board, gracefully chain contribution, hazard the executives is all under one shade of the development business. The mix of the record-based innovation has incited progress in weaving an organization of dealing with the entire framework has influenced decidedly through the employments of brilliant agreement-based frameworks and the IoT [5, 25]. Helpful in following standards of segments from the distribution centre to the site has demonstrated to dodge duplication of passages. The record framework gets refreshed by the bit by bit passage. The appropriated record framework likewise benefits the BIM framework to determine issues concerning the restricted cooperation and the information share. With the communitarian utilization of conveyed record framework in the BIM, the factor of trust among the various levels are to increment numerous folds regarding the development of a straightforward media including the task parties [26, 27]. This chronicle of the exchanges has prompted the affirmation of the reliable passage for instalments. The graceful chain

acquisition utilizing the keen agreements has started the robotized instalments, provenance following, contract organization, and a few others encouraging the contractual workers to build the degree of trust among the holders.

The significant issues that are confronted while actualizing the blockchain culture in the businesses are the sole function of the cost of establishment and execution utilizing IoT empowered gadgets; the other significant test is of the mindfulness and utilizing the framework by the various specialists in a venture that requires somewhat more opportunity to take up the cycle to learn. Even though the positive side of huge execution will prompt a superior framework yet will be a war-saw over the business that is managing the legally binding practices in the corporate field. The major influence will be the outsiders cause the blockchains target checking the inclusion of outsider intermediation [10, 28–31].

9 Research Method and Results

Developing a blockchain can be a brainchild to the thought or idea of introducing the blockchain advances in the fragmental construction industry, working on the process by developing it in the Remix environment and with the motto of mining the test ether coins having counted in terms that prove its validity, expiry terms the particular designed smart contract has in store the saved information that is necessary for a mutual agreement among the different stakeholders (Figs. 1, 2, 3 and 4).

The included inventory includes the different terms like the gas limit that indicates the range of the mining in the chain lying on the contract.

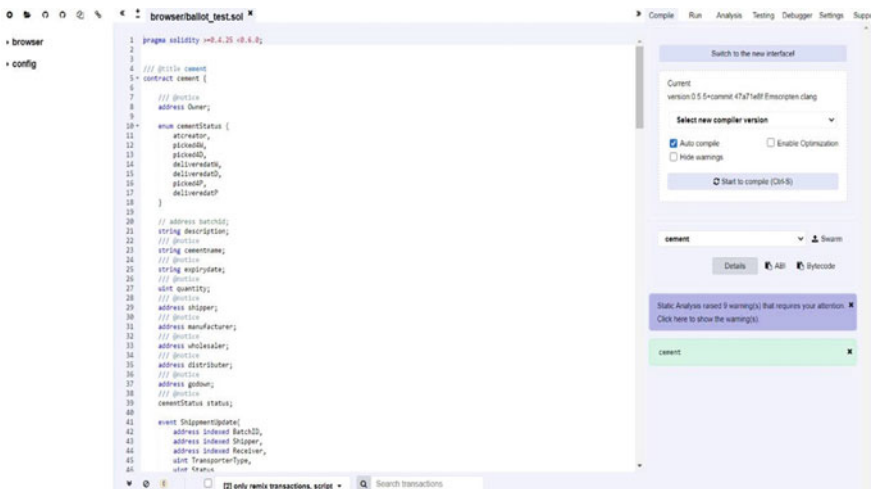


Fig. 1 Remix platform with code for starting smart contract



Fig. 2 Specifying the values for the contract

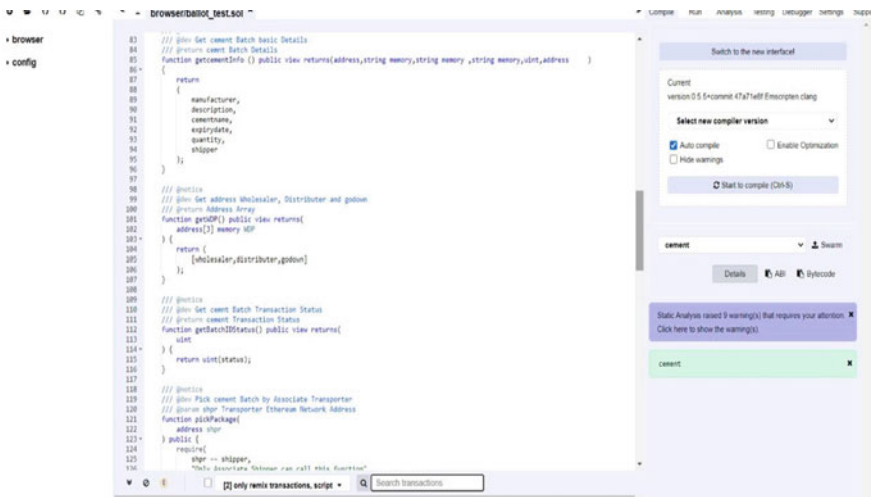


Fig. 3 Execution of the inventory mentioned

10 Conclusion

Through studying the various fronts of inculcating the blockchain technology in the construction industry has the potential to open a canopy to the biggest fragmented industry yet the major role is to let the professionals in the field understand and legitimize that the before stated adversaries are an escape route to the easy maintenance that only requires the knowledge of the different levels of the construction industry to be familiar and also willing to let share the network among the levels of the management at different steps. Thus, allowing the industry to minimize costs, reduce project

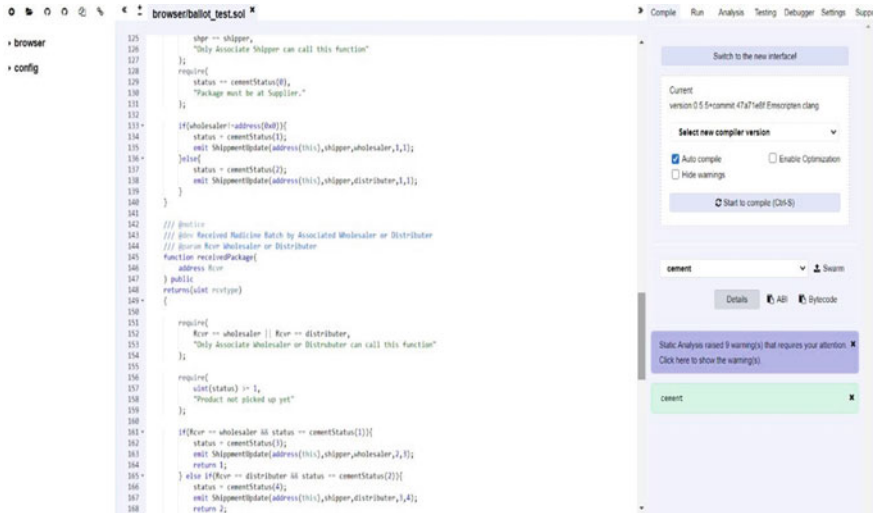


Fig. 4 Mining ether coins for the dry run of the contract

durations due to the correct estimation, and also quality assurance with blockchain in the supply chain procurement system settling payment disputes as well. Thus, with the view that the adoption of this technology will be beneficial with the advent of blockchain in the construction industry to gape the link between the customers and the stakeholders without the involvement of the third-party involvement will play a bigger role.

References

1. J. Garzik, J.C. Donnelly, Blockchain 101: an introduction to the future, in *Handbook of Blockchain, Digital Finance, and Inclusion*, vol 2 (Elsevier, Amsterdam, 2018), pp. 179–186
2. A. Koutsogiannis, N. Berntsen, Blockchain and construction: the how, why and when | BIM+. Chartered Institute of Building (2017). [Online]. Available: <https://www.bimplus.co.uk/people/blockchain-and-construction-how-why-and-when/>. Accessed: 30 July 2018
3. T. Aste, P. Tasca, T. Di Matteo, Blockchain technologies: the foreseeable impact on society and industry. *IEEE Comput. Soc.* 18–28 (2017)
4. M. Swam, *Blockchain Blueprint for a New Economy*, 1st edn. (O’Reilly Media, United States of America, 2015)
5. J. Wang, P. Wu, X. Wang, W. Shou, The outlook of blockchain technology for construction engineering management. *Front. Eng. Manage.* 4(1), 67–75 (2017)
6. S. Apte, N. Petrovsky, Will blockchain technology revolutionize excipient supply chain management? *J. Excipients Food Chem.* 7(3), 76–78 (2016)
7. M. Ramage, From BIM and blockchain in construction: what you need to know. Trimble Inc. (2018). [Online]. Available: <https://constructible.trimble.com/construction-industry/frombim-to-blockchain-in-construction-what-you-need-to-know>. Accessed: 30 July 2018
8. S. Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System (2008), p. 9. www.Bitcoin.Org.

9. J. Yli-Huomo, D. Ko, S. Choi, S. Park, K. Smolander, Where is current research on blockchain technology?—A systematic review. *PLoS ONE* **11**(10), 1–27 (2016)
10. J. Condos, W.H. Sorrell, S.L. Donegan, Blockchain Technology: Opportunities and Risks. Vermont, no. Vermont (2016), pp. 1–35
11. D. Drescher, *Blockchain Basics—A Non-technical Introduction in 25 Steps* (Springer Science and Business Media, New York, 2017), pp. 81–92
12. Ž Turk, R. Klinc, Potentials of blockchain technology for construction management. *Procedia Eng.* **196**(June), 638–645 (2017)
13. M.E. Peck, Do you need a blockchain? *Spectrum IEEE* 38–39, 60 (2017)
14. Z. Zheng, S. Xie, H. Dai, X. Chen, H. Wang, An overview of blockchain technology: architecture, consensus, and future trends, in *2017 IEEE International Congress on Big Data (BigData Congress)* (IEEE, 2017), pp. 557–564
15. S. Brakeville, B. Perepa, *Blockchain basics: Introduction to business ledgers* (2016)
16. I. Belle, The architecture, engineering and construction industry and blockchain technology, in *Digital Culture* (2017), pp. 279–284
17. V. Buterin, A next-generation smart contract and decentralized application platform (2014)
18. F. Glaser, Pervasive decentralization of digital infrastructures: a framework for blockchain enabled system and use case analysis (2017)
19. P.L. Seijas, S.J. Thompson, D. McAdams. Scripting smart contracts for distributed ledger technology. *IACR Cryptology ePrint Archive* (2016), p.1156
20. P.J. Taylor, T. Dargahi, A. Dehghantanha, R.M. Parizi, K.K.R. Choo, A systematic literature review of blockchain cyber security. *Digit. Commun. Netw.* (2019). <https://doi.org/10.1016/j.dcan.2019.01.005>
21. H.F. Atlam, G.B. Wills, Technical aspects of blockchain and IoT, in *Role of Blockchain Technology in IoT Applications* (Elsevier, Amsterdam, Netherlands, 2018), pp. 1–39
22. A.H. Mohsin, A.A. Zaidan, B.B. Zaidan, O.S. Albahri, A.S. Albahri, M.A. Alsaalem, K.I. Mohammed, Blockchain authentication of network applications: taxonomy, classification, capabilities, open challenges, motivations, recommendations and future directions. *Comput. Stand. Interfaces* (2018). <https://doi.org/10.1016/j.csi.2018.12.002>
23. M. Ali, J. Nelson, R. Shea, M.J. Freedman, Bootstrapping Trust in Distributed Systems with Blockchains (2016). Available online: https://www.usenix.org/system/files/login/articles/login_fall16_10_ali.pdf. Accessed on 21 Feb 2019
24. C. Eastman, P. Teicholz, R. Sacks, K. Liston, *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors* (Wiley, Hoboken, 2011)
25. A.R. Earls, *Blockchain Not a Panacea for Supply Chain Traceability, Transparency* (2016). Available at: <https://searchmanufacturingerp.techtarget.com/>. Accessed 21 Aug 2018
26. M. Mathews, *BIM and Blockchain* (2017). [Online]. Available: <https://iebc.co/bim-andblockchain/>. Accessed: 06 Aug 2018
27. K. Korpela, J. Hallikas, T. Dahlberg, Digital supply chain transformation toward blockchain integration, in *Proceedings of the 50th Hawaii International Conference on System Sciences* (2017)
28. S. Mandal, S. Biswas, V.E. Balas, R.N. Shaw, A. Ghosh, Motion prediction for autonomous vehicles from Lyft dataset using deep learning, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India (2020), pp. 768–773. <https://doi.org/10.1109/ICCCA49541.2020.9250790>
29. S. Mandal, V.E. Balas, R.N. Shaw, A. Ghosh, Prediction analysis of idiopathic pulmonary fibrosis progression from OSIC dataset, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India (2020), pp. 861–865. <https://doi.org/10.1109/GUCON48875.2020.9231239>

30. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering*, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh. Lecture Notes in Electrical Engineering, vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
31. Ethereum, Building a Smart Contract Using the Command Line (2018). Available at: <https://www.ethereum.org/greeter>. Accessed 02 Sept 2018

Dark Web—Onion Hidden Service Discovery and Crawling for Profiling Morphing, Unstructured Crime and Vulnerabilities Prediction



Romil Rawat, Anand Singh Rajawat, Vinod Mahor, Rabindra Nath Shaw,
and Ankush Ghosh

Abstract The mainstream use of the Internet and mobile technology has made it easier to reach a wide range of globalized resources. By having the numerous lawful design principles offered by Internet merchants, unregulated cyber terrorist and hacking markets by intruders actively working at online platform. Illicit web store and markets are capitalizing on the privacy and at globalized existence on the Internet, posing problems, hidden unsecure environment and obstacles for policing and security agencies. The proposed work offers a summary of the anonymous crime favored place at Dark Web and recent research on illegal Internet drug trafficking, human trafficking, terrorist funding and recruitment, money laundering, contract hacking, organ trafficking, cracked key distribution, killing contracts, forged passport, illegal post sharing and forum discussion, fraud and credit card selling, weapons order, and cybercrime and child sexual abuse markets in the Dark Web. The work presented here outlines the Tor network crawling procedure and evaluation of hidden links for analysis with the crawling of drug trafficking, Criminal activity-related signatures and posts put light on the negative side of the Dark Web platform and their services, techniques and methods use for data crawling, pattern recognition and behavior understanding of criminals followed by terrorist organizations, campaigning on social network platform using hidden identities for recruitment, fundraising and radicalization.

R. Rawat · A. S. Rajawat (✉)

Department of Computer Science Engineering, Shri Vaishnav Vidyapeeth Vishwavidyalaya,
Indore, India

V. Mahor

Department of Computer Science and Engineering, IPS College of Technology and Management,
Gwalior, MP, India

R. N. Shaw

Department of Electrical, Electronics and Communication Engineering, Galgotias University,
Greater Noida, India

e-mail: r.n.s@ieee.org

A. Ghosh

School of Engineering and Applied Sciences, The Neotia University, Sarisha, West Bengal, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

717

S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,

Lecture Notes in Electrical Engineering 756,

https://doi.org/10.1007/978-981-16-0749-3_57

Keywords Cybercrime · TOR (the onion router) · DarkWeb · Terrorist funding · Recruitment · Policing

1 Introduction

In the open access environment of web and anonymity provided at darkweb, it became easy for terrorists and criminals to share their propaganda using messages, videos, discussion forums, blogs, post, tweet, photos, code words to influence and lure the Internet users by showing opportunities, growth, money, success, short cut to make money, agents, recruiters, followers to manage the functioning of these online activities by channeling the events and connections, sometimes by continuous mind wash or by currency transfer for involvement in illicit activities. The money laundering, cryptocurrency is used to circulate the money around the criminal gangs and terrorist groups. The popularity of Internet is grown to high extent, and some countries fail to control the illicit media usage for criminal and terrorist activities.

Dark web is the hidden portion of Internet that is not indexed with any search engines and could only be accessed by URL address, having the length of 16 characters. The identifier accommodates the very first 80 bits of Public Key (RSA), and encoding is code in base32 value using SHA-1 Digest [1] and uses the top-level service of domain called Onion. The address keeps on changing that's why difficult for security agencies to track the online market, business, services and involved criminals. Mostly, illicit activities are done at unused IP address and traffic is diverted there. Around 60% information found on darkweb is illegal, and more than 3000 live online service [2] websites are working for dark web criminal business offering guns, pornography [3], drug, killing contract, organ trafficking, credit card, hacked movie service accounts and generating huge illegal revenue traveling across the world using cryptocurrency and money laundering [4] procedures in the form of donations, charity, trust, funds, etc.

Dark web platforms contain the same features and functionality as any online retail operation, adding ratings and reviews, shopping wallet carts [5], forums and carts but major variations remain. One is regulation of consistency. When buyers and sellers are private, every ranking system's integrity is questionable. Ratings are quickly distorted, and even long-track record dealers were reported to unexpectedly vanish with the cryptocurrencies of their clients.

Many e-Business and commerce companies provide screening services that containing the consumer funds and transactions information on hold until the ordered packets and product has been shipped. In the process do not expect a provider with a smile in the case of a disagreement. It is the understanding of the business persons and owners to draw it back. Each correspondence is encrypted and converted using secure algorithms, and PGP key [6] involves in whole process.

There is not the surety of product timely delivery at darkweb. The product needs to cross international boundaries by escaping from security agencies, and border security officials are cracking illicit imports. Deep Dot dark Web's platform [7] is full

of shops about customers convicted or imprisoned for attempted sales. There’s even information you wouldn’t be amazed to discover on the open platform, such as access to hard-to-find book full-text editions, political [8] news collections from popular blogs, and a Virginia Tech campus steam tunnel map. Intel Exchange will conduct anonymous exchanges on public affairs. There’s a number of informant pages, like WikiLeaks’ [9] dark network edition. Pirate Bay, BitTorrent’s den, frequently shut down [10, 11] by law enforcement, is alive and well. And Facebook’s online footprint is dark. Increasing numbers of cyber hackers, terrorist and criminals use the dark web platform using TOR which could not be accessed using normal search engines [12, 13]—like Bing, Yahoo, Google, like Tor browsers.

The web structure and their structure have been shown consisting markets and sites access information in Fig. 1.

It becomes difficult for cyber police to identify the crime committed on dark web platform, and continuous incidents have been registered like malware hacked system information containing financial audits. But still police designing cyber operations, where official’s works in covert mode at dark web and deals like criminals to track the illegal activities, but it may take years to found the origin and sources of illicit activities.

Psychological warfare informatics is alluded to as the use of cutting edge data combination, examination strategies and systems to gain measure, coordinate, oversee and dissect the variety of the data identified with illegal intimidation for



Fig. 1 Internet classification into different web platforms—surface web, dark web, deep web [25, 26]

global/public security applications. The strategy is obtained from the trains, for example, informatics, arithmetic, science, insights, sociologies, public approach, and etymology. The exploration shows illegal intimidation includes a colossal measure of data from various sources, dialects, information types, data combination and investigation, for example, text mining, information mining [7], language interpretation, information mix, video and picture preparing assists with recognizing and forestall illegal intimidation.

The purpose of cyber criminals at darkweb of misrepresentation and robbery is applicable at both the public and global levels, since crooks may escape by utilizing bogus characters, and the dealers can likewise enter the nation by holding counterfeit visas or travel papers. Web misrepresentation, network hacking, interruption, unlawful exchanging, scorn wrongdoings, infection spreading, digital sexual entertainment, digital security, robbery of the classified data what's more, digital illegal intimidation, opiates dealing and psychological oppression have no limits what's more, are a security concern around the world.

With the advances in data innovation hoodlums are utilizing the internet to submit various digital violations, developing patterns of complex dispersed and Internet processing raise significant inquiries regarding data security and protection. Digital foundations are profoundly defenseless against interruptions and different dangers. Actual gadgets, for example, sensors and finders, are most certainly not adequate for checking and insurance [12] of these foundations; henceforth, there is a requirement for additional advanced IT that can show ordinary practices and distinguish unusual ones. These digital guard frameworks should be adaptable, versatile and hearty, and ready to recognize a wide assortment of dangers and settle on clever continuous choices.

With the movement and measure of digital assaults, human intercession is basically not adequate for ideal assault investigation and suitable reaction. The truth of the matter is that the most organization-driven digital assaults are done by keen specialists, for example, PC worms and infections; thus, battling them with clever semi-independent specialists that can identify, assess, and react to digital assaults has become a prerequisite. These purported PC-produced powers should have the option to deal with the whole cycle of assault reaction in an ideal way, for example to finish up what kind of besides digital interruptions [3] are not confined. They are a worldwide danger that presents danger to any PC framework on the planet at a developing rate. There were times when just instructed subject matter expert could perpetrate digital violations; however, today with the development of the Internet, nearly anybody has admittance to the information and instruments for perpetrating these wrongdoings. Regular fixed calculations (hard-wired rationale on dynamic level) have been ineffectual against battling progressively developing digital assaults. This is the reason we need imaginative methodologies, for example, applying techniques for artificial intelligence (AI) that give adaptability and learning capacity to programming which will help people in battling digital violations [4, 5] Man-made intelligence offers this and different conceivable outcomes. Various nature-motivated registering strategies for AI (e.g., computational intelligence, neural networks, intelligent agents, artificial immune frameworks, machine learning, data mining, pattern recognition,

fuzzy logic, heuristics, and so forth) have been progressively assuming a significant part in digital wrongdoing discovery and counteraction. Man-made intelligence empowers us to plan autonomic figuring arrangements fit for adjusting to their setting of utilization, utilizing the techniques for self-administration, self-tuning, self-setup, self-conclusion, and self healing. With regard to the fate of data security, AI procedures appear to be exceptionally encouraging territory of exploration that centers on improving the safety efforts for the Internet.

1.1 Research Groups Continues Working Toward Advance Technology for Generating the Protective Shield Against Cyber Crime

1. Advanced securing feature enabling in operating system and distributed environment.
2. Data structure and algorithms analysis to forecast loop holes.
3. Use of machine learning, artificial intelligence, deep learning, big data for analysis of content and working procedures of criminals.
4. Strong analysis and model designing for software engineering, business data analytics, digital marketing.
5. Guiding and educating police, security agencies toward advanced tactics of cyber security and crime, cyber terrorism, forensic, online social network security.
6. Removing of malfunctioned and generating of best design for computer architecture, parallel processing and organization.
7. Development of strong security rules in theory of computation, compiler, logical mathematics, and discrete structure.
8. Compulsory and automatic enabling of security algorithm when any malicious signature appears in embedded system, IoT, hardware programming, robotics.

The rest of the paper is organized as follows. Section 2 covers illegal activities conduction at TOR network. Anonymous structure of data transfer at dark Web [Tor-The Onion Router] will be carried out presented in Sect. 3, Sect. 4 describes terrorist funding, process and source, Sect. 5 gives detail about positive side of darkweb, Sect. 6 works for organization of the web, Sect. 7 provides related work, Sect. 8 shows Dark web service architectures, Sect. 9 concludes the paper with future directions.

2 Illegal Activities Conduction at TOR Network

Cyber criminals have created dark web online stores and market [14] for the purchase of illegal products named as: Astrid, Avengers, Backopy, Cirrus, Defcon, Digitalink, Dread Pirate Roberts (Ross William Ulbricht), Flush (Curtis Green), Silk Road,

Libertas, Lucskyhigher, Tony76, Dream market, Empire market, Nightmare market, agartha market, core market, Cryptonia market, cannazon, CGMC, AlphaBay Market

The list is not exhaustive but selected based on the availability and popularity.

Security agencies and police tried to trace the illegal dark web market and succeeded several times, but new vendor site with different name appears on darkweb and the information is secretly transformed to clients for business.

2.1 Dark Web Crimes

1. Drugs Trafficking

Different flavors of drug are available at online store and could be accessible throughout the world working anonymously, most famous one the Silk Road market.

2. Illegal Business of Looted Antiquities and Artifacts (Digital and Ancient) on the Dark Web

Illegally stolen antiquities and artifacts are made available on the darknet platform and once buyer shows interest, the business starts and the artifacts need to travel across border.

3. Human Trafficking

Number of incidents happen when human goes on missing, some cases were registered toward investigation by security agencies and but others goes unnoticed and that is how human trafficking business works, they searches for lone, poor, needy people shows perks and makes them invisible for pornography, illegal activity conduction-type activities, and Fig. 2 displays the purpose.

4. Organ Trafficking

The organ could not be stored for longer time and has limitation of same genetic order for transmission, the needy have to wait for longer time to get it legally by government hospitals that's why the organ trafficking business came the agents, donor, needy person, receiver, surgery doctor create communication channels by sharing secret techniques, so that only person involved in the business can work secretly, but every illegal activities come with fraud, the donor and racier has to travel to multiple countries to make the deal, supply amount in the form for cryptocurrency and other ways, so that it could not be noticed by security agencies.

1. Contract hacking, Illegal weapon selling Market and attack generation

The contract cyber criminals and tools are available for conducting malicious activities. The all weapons are made to sell at the darkweb platform, but the issues faced by criminals are at delivery, because escaping at multiple layer by security personal is difficult.

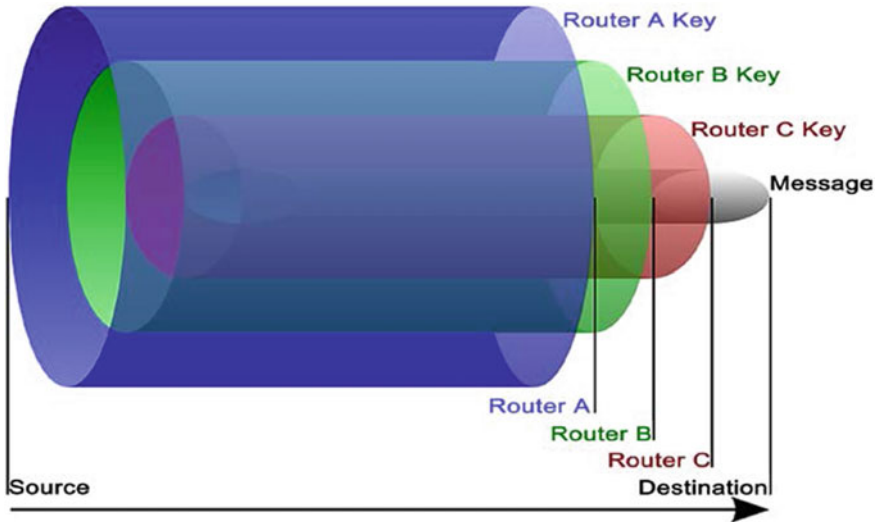


Fig. 3 Onion routing architecture [23]

3 Anonymous Structure of Data Transfer at DarkWeb [TOR—The Onion Router]

Once downloading and installing of the TOR browser is done, Tor directory searches for three anonymous nodes (router) for traffic routing called as entry node, relay node and exit node. Each node contains the specific key for opening the packet, and after that, it processes the information and passes to next node, just like Onion layers. Even the consecutive node does not have address of each other, and the process keeps the secure transmission. Figure 3 shows the layer architecture with a key for opening the packet.

4 Terrorist Funding, Process and Source

The terrorism has moved toward cyber platform and achieved the global recognition, they are most active in social network and at hidden platform, they are recruiting criminals and cyber hackers or computer operators to design and share their propaganda, through which fund can be generated and transferred using money laundering to intended location. The below diagram depicts the process followed by criminals and terrorist for rounding the money and generating fund. The recruitment for terrorist shown in Fig. 4.

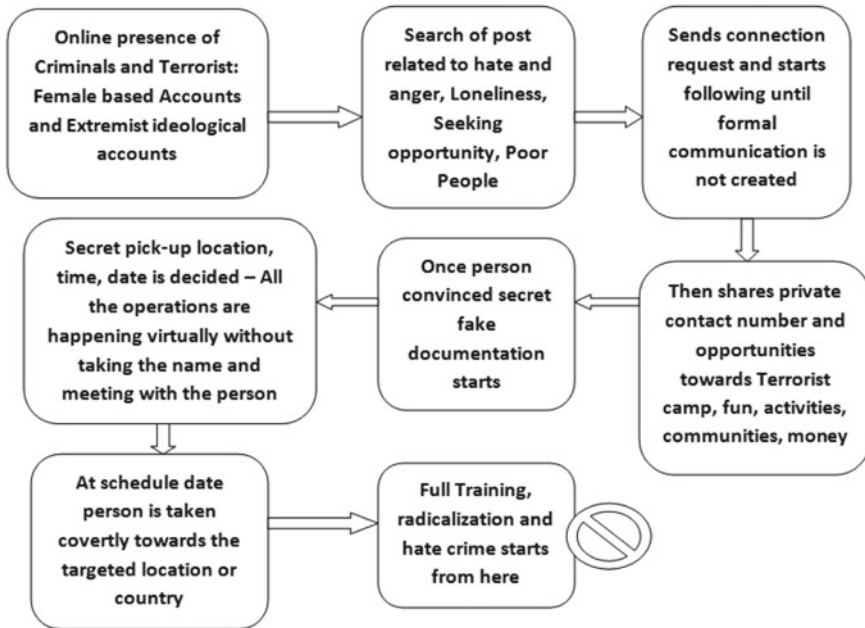


Fig. 4 Terrorist recruitment activities

5 Positive Side of Darkweb

Although having several criminals activities, darkweb could also be used useful and safe works.

1. It is used by journalists, whistle blowers to show and publish the negative and conspiracy-related information about government, organization and politician, which should be known by people, but have threat of life for the publishers.
2. Lots of free books and digital libraries are there for study.
3. Survivors put their stories about uncomfortable experiences.
4. Dark web market products like weapon could also be sued for personal safety.
5. It shows access to outer world by dark web Internet, when any country imposes restriction for their citizens or blocks the Internet contents.
6. Some drugs used for medicinal purpose and could be purchased for the treatment.
7. Users are anonymous at darkweb, if system security is enabled, could be safe place for internet browsing by having unnoticed.
8. Darkweb provides freedom of expression and privacy, when counted for good.
9. Free from unnecessary advertisement an pop-ups.
10. Police and security agencies work as undercover agents to track criminal activities

6 Organization of the Web

Surface web is called as public data, and only around 4% pages are linked to popular and normal search engines [15, 16] like Twitter, Facebook, Reddit, Wikipedia, Instagram, news sites, educational sites, etc.

Deep Web is called as private data, and around 90% data contain private mails, chats, video, photos, bank accounts details, cloud personal data and could be accessed by authorized login and password.

Dark web is called as anonymous data, and around 6% data live here, nobody is allowed to access it by normal procedures, here special consideration and browser is required. All website extensions are marked as Onion. Even Facebook having the Onion—the darknet version—facebookcore wwvi.onion [17]. The most popular threat among business rivals are the risk variables devaluating a company that may involve deteriorating brand image, reputational conspiracy, damage and loss by competitor ruining a business that may include RDoS, DDoS [18, 19] attacks, ransom ware, malware impacting business services and operations. Fraudging a corporation that may include IP fraud, virus [20] transfer or spying that brings effect on company’s capacity to perform or direct and huge financial losses. Figure 5 shows the types of activities done at different architectures.

	Surface Web	Deep Web	Dark Web	Darknet
Description	Content that search engine can find	Content that search engine cannot find	Content that are hidden intentionally	-
Known as	Visible web, Indexed web	Invisible web, Hidden web, Deep web	-	Underbelly of internet
Constitutes	Web	Web	Web	Web
Contents	Legal	Legal + Illegal	Illegal	Network
Information found	4%	96%	-	-
Browsers	Google chrome, Mozilla firefox, opera	-	TOR browser	Freenet, Tor GNUnet, I2P, OneSwarm, RetroShare

Fig. 5 Web activities [28]

7 Related Work

Kaur et al. [2] an summary of numerous forms of threats, vulnerabilities, and ransomware is also presented. Different forms of illegal acts and events are addressed at the dark web platform and defined the preventive measures.

Godawatte et al. [4] deep-end study would explore the dark network connectivity anonymously and how law enforcement authorities exponentially monitor people with terror habits and actions. The paper also addresses dark websites where people may snatch dark network jihadist resources and anonymous markets, including security measures.

Zulkarnine et al. [5] shown the frequent websites working at dark web and acting as catalyst agent by offering the knowledge base, support and utilities required to create Tor secret services and onion websites.

Akyildiz et al. [6] Deep Web and Dark Web definitions and partnerships were analyzed. Deep Web and Dark Web's operating processes investigated, the data collected by these media analyzed, and the data collected assessed in terms of information protection. Deep Network and Dark Web worlds in cyber actions' illegal material are graded.

Nazah et al. [7] this SLR provide a detailed overview of Dark Net crime risks, technological and investigative problems of anonymous network systems, identification mechanisms, techniques, resources and tactics used to identify crimes and offenders on the Dark Web. Cyber offenders become quicker toward enforced methods of identifying them inside.

Dark Web, Park [8] propose a main concept focused on the present state of those crimes and the Deep Network Crime Knowledge Visual System. Deep network status is evaluated and Java visualizes details. The software is intended to enable more effective crime management and tracking in unknown network like deep web, torrent, etc.

Proposed the advance system (advance terrorist detection system) for detecting the terrorist activities [5] is much better than the previous one called TDS (terrorist detection system) and has achieved best score in illicit activity detection by automatically tuned tracking of terrorist recruitment and radicalization activities.

The approach for tracing the techniques of catching the online markets offering illegal services like drug supply, arms trafficking, pornographic contents, human trafficking, organ trafficking [15], etc.

The identification, tracing and searching of terrorist activities [12] includes money laundering, cyptocurrency operations, fund raising, and terrorist-related content sharing.

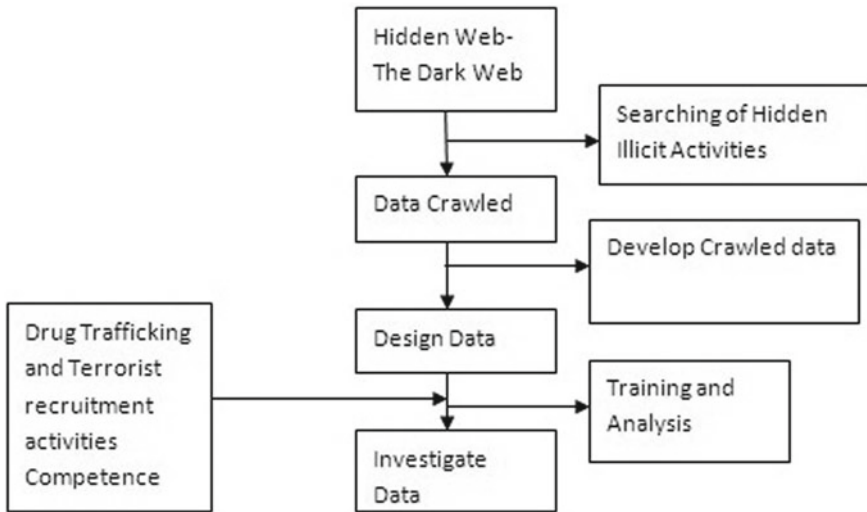


Fig. 6 Methodology of dark web mining

8 DarkWeb Service Architectures

8.1 DarkWeb Mining

The mining of darkweb platform is required to find the patterns and crawl data for the purpose of analysis. Figure 6 depicts the process of drug trafficking and terrorist recruitment activities competence. Each block represents the process and node for further data processing.

8.2 The Tasks and Interaction of Tools in the Process of Hidden Service Discovery

The Onion link extractions are done to identify the hidden contents of dark web as depicted in Fig. 7. The node represented by numbering represents the algorithm steps and task for processing by preceding data transfer.

8.3 DarkWeb Tools and Threats

The services and communities using the dark web tools for criminals activities shown in Fig. 8. The main purpose of cyber criminals to scan login credentials, credit card

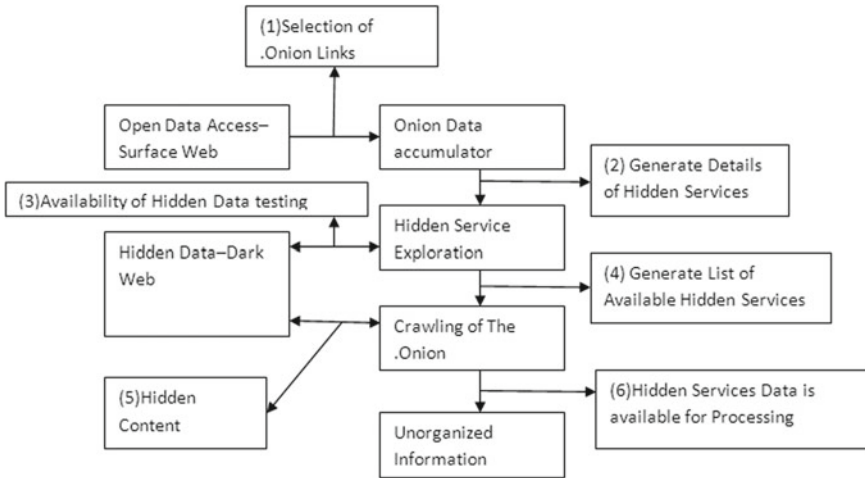


Fig. 7 Hidden service discovery

DARK WEB TOOLS	SERVICES	PRESENT ENTERPRISE RISK
Infection or attacks, including malware, DDoS, RDoS, and botnets	Ransomware, Terrorist groups	Oil and Gas Automated Industries
Espionage, Funding, radicalization.	Terrorist groups,	Political Groups
Remote access Trojans (RATs), Exploits, keyloggers	Extremism	Whistleblower, hoaxes.
Support services such as tutorials and Manuals	Terrorist groups	Banks and financial organizations, Online recruitment and radicalization
Credentials	Cyber Criminals	Social Network Fraud
Phishing	Online Harassment	Contract killing and extortion
Refunds	Identity Theft scams	Eavesdropping attack
Customer and Financial data	Cyber stalking	Arms and Violence
Operational data	Invasion of privacy	Cyberpolicing
Drug Market	Cryptocurrency	Money laundering, Petroleum Industry, Ministry
Organ Trade	Human Trafficking, child pornography	Unauthorized arm handling and healthcare hacking.

Fig. 8 Dark web tools and services that present enterprise risk [27]

numbers, IP addresses, open ports, security patches in system so that malicious threat could be easily triggered. Security agencies and organizations are continuously working to find the vulnerabilities and illegal activities, and it is done by darkweb scanners, site monitors, web scrapers.

DARK WEB MONITORING TOOLS FOR NETWORK ADMINS	TASK	LIMITATION	FEATURE
Echosec Beacon	Dark Web scanner	impose unique limits on search parameters	To find stolen credential
SpyCloud ATO Prevention	Online service protection	Has limit towards control security	Resetting of exposed passwords
Digital Shadows SearchLight	Corporate Brand ,Image, worth and reputation Protection .	Access to Expertise	Detects data loss
WhatsUp Gold	network monitoring system	Device limitation	Network monitoring and alerting
DigitalStakeout Scout	TIS-threat intelligence service	Cyber physical System is susceptible to attack.	provides Privacy Policy
Alert Logic Dark Web Scanner	Information Protective system	limits on scan jobs	account takeover
DarkOwl Vision	malicious data site indexing	Issue in capcha and verification	Identification of Stolen data
ACID Cyber Intelligence	Cyber Intelligence Tool for Social Network Platform	information sharing and the analysis tools	Collects Vulnerability and illicit behavior at social networks, Chat portal, forums.
Dashlane Business	password access protection, secure file access storage,	Cyber physical System is susceptible to attack.	account protection service

Fig. 9 Dark web monitoring tools [29]

8.4 Darkweb Scanners

See Fig. 9.

8.5 Dark Web Data Searching Platforms and Engines

The darkweb content could be explored using several special search engines with their applications presented in Fig. 10.

8.6 Dark Web Warning

It is likely good to secure the system before accessing the dark web, install antimalware, antispysware, antivirus software and use of dark web vulnerability security tools [21–24] like ExpressVPN, NordVPN, Surfshark, and IPVanish so that your presence could go unidentifiable and not traced by attackers of darkweb environment and several web browsers are there for darkweb surfing anonymously presented here in Fig. 11.

DARK WEB SEARCH ENGINE	APPLICATION	LINK
Hidden Wiki	Wikipedia Containing links to several sites working on darkweb	http://zqktlwuaavvvqqt4ybvgvi7tyo4hj15xgfuvpdf6oijycgwbqym2qad.onion/wiki/index.php/Main_Page
DuckDuckGo	Generate list of .onion websites	https://3g2upl4pq6kufc4m.onion/
Candle	through the dark web, figuratively speaking	http://gjobqjj7wyczbqie.onion/
Not Evil	Shows deep contents	http://hss3uro2hsxfogfq.onion/
SearX	Regular engine for searching	http://ulm6sryqaiiefld.onion/
Facebook	Dark web facebook version.	https://www.facebookcorewwi.onion/
Bitcoin and other cryptocurrencies	cryptocurrencies wallet - Smartmixer ,Onion, Green Address	http://ow24et3tetp6vmk.onion/
BBC Tor Mirror	access to Access of block foreign news sites	http://bbcnewsv2vjptsuy.onion/
Tor Metric	Tor Project Information	http://rougmnvswfsm4dq.onion/
Hidden Answers	Darkweb version of Quora and Reddit.	http://answerstedhctbek.onion/
SecureDrop	whistleblowers ,journalists ,Spying communication channel	http://secdrop5wyphb5x.onion/
ProtonMail	Service of Swiss encrypted and secured email	https://protonirockerxow.onion/
SecMail	Service of Swiss encrypted and secured email	http://secmailw453j7piv.onion/
Mailpile	Service of Swiss encrypted and secured email	http://clgs64523yi2bkhx.onion/
Comic Book Library	Open Comic Book Library Access	http://r6rfy5zlfbsiym.onion/
Deep Web Radio	Open music Library Access	http://76qugh5bey5gum7l.onion/
Ahnia	Traditional Like search Engine	msydgstlz2kzerdg.onion
Kilos	Traditional Like search Engine	dnmugu4755642434.onion
HayStak	Traditional Like search Engine	haystakvxd7wbk5.onion
Abiko	Traditional Like search Engine	abikogailmonxlz1.onion

Fig. 10 Dark web data searching platforms and engines [14, 30]

BROWSER NAME	ADVANTAGE	LIMITATION	FEATURES	GROUPS
The Tor Browser	Anonymous browser.	speed	Complex Data encryption	Extremism , Cyber attackers
Subgraph OS	Kernal Hardening, Metaproxy, and File System Encryption	security hardening	reduce the attack surface	Cyber Criminals, Extremism
Firefox	Tor Network	Too much memory consumption	Best platform for anonymous work	Whistle Blower, Cyber attackers
Waterfox	Firefox Like browser, Waterfox	Too much memory consumption	reduce the attack surface	Cyber attackers, contract cyber criminals
Whonix	Open browsing	Problem in metadata	data directory is available after reboot	Hackers, Government spies, Cyber attackers

Fig. 11 Web browsers for anonymous web surfing [31, 32]

9 Conclusions

Darkweb is also called Invisible Internet Project (I2P), and the Onion Routing Project (TOR) is a portion of the web environment designed in USA naval science secret laboratory by DARPA-Defense Advanced Research Project Agency for the anonymous transfer of military information but later it is open for all now the darkweb became hub for criminal and terrorist activities, because the very less fear to be identified by the security agencies due to the features of unavailability of indexing contents,

it makes dedicated space. The criminals mostly use unused IP address and duns the network traffic for business operations. The policing and cyber security agencies having the only chance to be covert and the part of cyber criminals process until they identified the authentic source, name, place and information based on which they can capture the criminals. In the proposed work, a framework of darkweb environment and criminals activities is presented. The series and range of attacks are working on the platform and even created as per need, if one is traced by police other comes in existence.

Future Directions

The MEMEX project has been launched and still in progress to index all. Onion websites, so that it could be listed at the search and analyzed by security forces and action against cruel and money-minded criminal communities business could be traced and stopped.

References

1. R. Liggett, J.R. Lee, A.L. Roddy, M.A. Wallin, The Dark Web as a Platform for Crime: An Exploration of Illicit Drug, Firearm, CSAM, and Cybercrime Markets, in *The Palgrave Handbook of International Cybercrime and Cyberdeviance*, ed. by T. Holt, A. Bossler (Palgrave Macmillan, Cham, 2020). https://doi.org/10.1007/978-3-319-78440-3_17
2. S. Kaur, S. Randhawa, Dark Web: A Web of Crimes. *Wireless Pers. Commun.* **112**, 2131–2158 (2020). <https://doi.org/10.1007/s11277-020-07143-2>
3. C.J. Didonato, S.E. Ingraham, J.R. Mendell, T.W. Prior, S. Lenard, R.T. Moxley III, J. Florence, A.H.M. Burghes, Deletion and conversion in spinal muscular atrophy patients: is there a relationship to severity? **41**(2), 230–237 (1997). <https://doi.org/10.1002/ana.410410214>
4. K. Godawatte, M. Raza, M. Murtaza, A. Saeed, Dark Web Along With The Dark Web Marketing And Surveillance, in *2019 20th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)*, Gold Coast, Australia, 2019, pp. 483–485. <https://doi.org/10.1109/pdcata46702.2019.00095>
5. A.T. Zulkarnine, R. Frank, B. Monk, J. Mitchell, G. Davies, Surfacing collaborated networks in dark web to find illicit and criminal content, in *2016 IEEE Conference on Intelligence and Security Informatics (ISI)* (2016). <https://doi.org/10.1109/isi.2016.7745452>
6. O. Akyildiz, [*IEEE 2018 6th International Symposium on Digital Forensic and Security (ISDFS)—Antalya (2018.3.22-2018.3.25)*] *2018 6th International Symposium on Digital Forensic and Security (ISDFS)—Information analysis and cyber crimes in Deep Web & Dark Web*, (2018), pp. 1–6. <https://doi.org/10.1109/isdfs.2018.8355358>
7. S. Nazah, S. Huda, J. Abawajy, M.M. Hassan, Evolution of Dark Web threat analysis and detection: a systematic approach. *IEEE Access* **8**, 171796–171819 (2020). <https://doi.org/10.1109/access.2020.3024198>
8. W. Park, A study on analytical visualization of deep web, in *2020 22nd International Conference on Advanced Communication Technology (ICACT)* (2020). <https://doi.org/10.23919/icaact48636.2020.9061283>
9. C. Gokhale, Network analysis of dark web traffic through the geo-location of South African IP address space, in *Smart Cities Performability, Cognition, & Security. EAI/Springer Innovations in Communication and Computing*, ed. by F. Al-Turjman (Springer, Cham, 2020). https://doi.org/10.1007/978-3-030-14718-1_10
10. K. Williams, Untangling the dark web: taking on the human sex trafficking industry. *IEEE Women Eng. Mag.* **7**(2), 23–26 (2013). <https://doi.org/10.1109/MWIE.2013.2280387>

11. C. Gokhale, O.O. Olugbara, Dark web traffic analysis of cybersecurity threats through South African Internet protocol address space. *SN Comput. Sci.* **1**, 273 (2020). <https://doi.org/10.1007/s42979-020-00292-y>
12. R. Liggett, J.R. Lee, A.L. Roddy, M.A. Wallin, The dark web as a platform for crime: an exploration of illicit drug, firearm, CSAM, and cybercrime markets, in *The Palgrave Handbook of International Cybercrime and Cyberdeviance*, ed. by T. Holt, A. Bossler (Palgrave Macmillan, Cham, 2019). https://doi.org/10.1007/978-3-319-90307-1_17-1
13. E. Mead, N. Agarwal, Surface web vs deep web vs dark web, in *Encyclopedia of Big Data*, ed. by L.A. Schintler, C.L. McNeely (Springer, Cham, 2020). https://doi.org/10.1007/978-3-319-32001-4_461-1
14. R.W. Gehl, Archives for the dark web: a field guide for study, in *Research Methods for the Digital Humanities*, ed. by Levenberg., T. Neilson, D. Rheams (Palgrave Macmillan, Cham, 2018). https://doi.org/10.1007/978-3-319-96713-4_3
15. H.S. Kassab, J.D. Rosen, Illicit markets and the internet age, in *Illicit Markets, Organized Crime, and Global Security* (Palgrave Macmillan, Cham, 2019). https://doi.org/10.1007/978-3-319-90635-5_8
16. M. Todorof, FinTech on the Dark Web: the rise of cryptos. *ERA Forum* **20**, 1–20 (2019). <https://doi.org/10.1007/s12027-019-00556->
17. M. Pannu, I. Kay, D. Harris, Using dark web crawler to uncover suspicious and malicious websites, in *Advances in Human Factors in Cybersecurity. AHFE 2018. Advances in Intelligent Systems and Computing*, ed. by T. Ahram, D. Nicholson, vol. 782 (Springer, Cham, 2019). https://doi.org/10.1007/978-3-319-94782-2_11
18. P. Ahonen, et al., Dark scenarios, in *Safeguards in a World of Ambient Intelligence. The International Library of Ethics, Law and Technology*, D. Wright, M. Friedewald, Y. Punie, S. Gutwirth, E. Vildjiounaite, vol 1 (Springer, Dordrecht, 2008). https://doi.org/10.1007/978-1-4020-6662-7_3
19. H. Chen, Dark web research overview, in *Dark Web. Integrated Series in Information Systems*, vol 30 (Springer, New York, 2012). https://doi.org/10.1007/978-1-4614-1557-2_1
20. J. Weber, E.W. Kruisbergen, Criminal markets: the dark web, money laundering and counter-strategies. *An overview of the 10th Research Conference on Organized Crime. Trends Organ Crim* **22**, 346–356 (2019). <https://doi.org/10.1007/s12117-019-09365-8>
21. S. Mandal, S. Biswas, V.E. Balas, R.N. Shaw, A. Ghosh, Motion prediction for autonomous vehicles from Lyft dataset using deep learning, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, 2020, pp. 768–773. <https://doi.org/10.1109/iccca49541.2020.9250790>
22. S. Mandal, V.E. Balas, R.N. Shaw, A. Ghosh, Prediction analysis of idiopathic pulmonary fibrosis progression from OSIC dataset, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, 2020, pp. 861–865. <https://doi.org/10.1109/gucon48875.2020.9231239>
23. M. Kumar, V.M. Shenbagaraman, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering*, ed. by M.N. Favorskaya, S. Mekhilef, R.K. Pandey, N. Singh (Springer, Berlin, 2020), pp. 765–773. ISBN 978-981-15-4691-4
24. J.N. Pelton, I.B. Singh, Coping with the dark web, cyber-criminals and techno-terrorists in a smart city, in *Smart Cities of Today and Tomorrow* (Copernicus, Cham, 2019). https://doi.org/10.1007/978-3-319-95822-4_11
25. A. Chetry, U. Sharma, Dark web Activity on Tor—investigation challenges and retrieval of memory artifacts, in *International Conference on Innovative Computing and Communications. Advances in Intelligent Systems and Computing*, ed. by D. Gupta, A. Khanna, S. Bhattacharyya, A.E. Hassanien, S. Anand, A. Jaiswal, vol. 1165 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-5113-0_80
26. F.M. De Sanctis, Dark web: deterring cybercrimes and cyber-attacks, in *Technology-Enhanced Methods of Money Laundering* (Springer, Cham, 2019). https://doi.org/10.1007/978-3-030-18330-1_2

27. H. Chen, Exploring extremism and terrorism on the web: the dark web project, in *Intelligence and Security Informatics. PAISI 2007. Lecture Notes in Computer Science*, ed. by C.C. Yang, et al., vol. 4430 (Springer, Berlin, 2007). https://doi.org/10.1007/978-3-540-71549-8_1
28. C. Iliou, G. Kalpakis, T. Tsirikas et al., Hybrid focused crawling on the surface and the dark web. *EURASIP J. Info. Secur.* **2017**, 11 (2017). <https://doi.org/10.1186/s13635-017-0064-5>
29. S. Abraham, P.S. Lal, Spatio-temporal similarity of web user session trajectories and applications in dark web research, in *Intelligence and Security Informatics. PAISI 2011. Lecture Notes in Computer Science*, ed. by M. Chau, G.A. Wang, X. Zheng, H. Chen, D. Zeng, W. Mao, vol. 6749 (Springer, Berlin, 2011). https://doi.org/10.1007/978-3-642-22039-5_1
30. F. Thomaz, C. Salge, E. Karahanna et al., Learning from the Dark Web: leveraging conversational agents in the era of hyper-privacy to enhance marketing. *J. Acad. Mark. Sci.* **48**, 43–63 (2020). <https://doi.org/10.1007/s11747-019-00704-3>
31. T. Singletary, Dark web and the rise of underground networks, in *Evolution of Cyber Technologies and Operations to 2035. Advances in Information Security*, ed. by M. Blowers, vol. 63. Springer, Cham, 2015). https://doi.org/10.1007/978-3-319-23585-1_8
32. I. Rola, D. Balzarotti, I. Santos, The onions have eyes: a comprehensive structure and privacy analysis of tor hidden services, in *Proceedings of the 26th International Conference on World Wide Web* (ACM, Perth, Australia), pp. 1251–1260 (2017). <https://doi.org/10.1145/3038912.3052657>

Suspicious Big Text Data Analysis for Prediction—On Darkweb User Activity Using Computational Intelligence Model



Anand Singh Rajawat, Romil Rawat, Vinod Mahor, Rabindra Nath Shaw, and Ankush Ghosh

Abstract In this paper, we work on suspicious big text data analysis technique for prediction of terrorism activity like financial fraud, money laundering, recruitment, radicalization, fundraising, violent and illegal post and video sharing on darkweb environment also called as cosmic web due to hidden content attributes. The consequent activity prognosis (CAP) is required for minimizing the risk associated with cyber information and personal security compromise for collectively data analysis referred as big data. The cyberterrorist and criminal hackers generated denial of service attack (DoS), distributed DoS attack (DDoS) and ransom-related DoS attack (RDoS) attack thereby overloading the server and increasing and blocking the server execution. The cyber threats and activities could only be reducing the execution time of activities marked suspicious and not safe. The propose model is based on computational intelligence technique using MapReduce technique, by classifying the malicious patterns found in big data sets collected from authentic channels and designed using machine learning supporting languages to implement the enhanced model and magnify the existing Intelligent techniques of computation with evaluated parameters. The work is highly adaptable for analysis and outline terrorist and criminal activities and would be beneficial for cyber police and security agencies.

Keywords Semi-supervised support vector machine · Computational intelligence · Big data · Terrorism activity · Darkweb

A. S. Rajawat · R. Rawat

Department of Computer Science Engineering, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore, India

V. Mahor

Department of Computer Science and Engineering, Gwalior Engineering College, Gwalior, India

R. N. Shaw

Department of Electrical, Electronics and Communication Engineering, Galgotias University, Greater Noida, India

e-mail: r.n.s@ieee.org

A. Ghosh (✉)

School of Engineering and Applied Sciences, The Neotia University, Sarisha, West Bengal, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

735

S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,

Lecture Notes in Electrical Engineering 756,

https://doi.org/10.1007/978-981-16-0749-3_58

1 Introduction

Computational intelligence (CI) is covenanted by construction of classifications which is analyzed by information deprived from the essential of unambiguous programming. The computational intelligence (CI) algorithms can be categorized into support vector machine and semi-supervise support vector machine. Algorithms are expressed by new framework for substantial improvement in performance by using non-iterative algorithms. Semi-supervised support vector machine (S3VM) is optimization technique proposed by Olivier Chapelle et al. [1]. It is a method of support vector machines (SVMs) that can grip moderately labeled data sets which have logically remained energetically active subject. S3VM is majorly used for determining the difficulties in normal SVM, difficulty here is to add on unrecognized labels in a best way in the form of source as well changing values. It has to be accomplished by destroying the limit of unidentified (tagged) text data in any of the terrorism activity—on darkweb, although very specific studies have come to a conclusion that boundary which navigates thoroughly completing the compact text data density and elevating the taking part of the recognized labels in the given space. The complete objective of computational intelligence is to repeatedly construct models from information deprived and which need full monotonous and time unbearable hominid contribution. However, this objective has not been accomplished. One of the problems is that learning algorithms require parameter tuning in instruction to acclimatize them to the specifics of a training set. The machine learning community regulates how to improve hyperparameters professionally and deprive in overfitting. For example, S3VM classification necessitates an initial learning phase in which the training data are used to regulate the classification parameters.

There are numerous nearby S3VM algorithms and active learning approaches, which is insufficient to address the limitation in modification phase of learning which is life-threatening to accomplish high-quality consequences. Since exertion of the modification parameters is optimally, through the accumulative obtainability of inexpensive big amount of text (suspicious or terrorism activity—on darkweb) this has become additional practical. Cloud computing made enormous parameter investigation so easy that a novel problem may arise. The ultimate accuracy of the model on assessment data can be reduced by consecutively applying too numerous tests and over fitting to the validation data. In this research work, the design new MapReduce model is derived using computational intelligence (CI) for proficiently perform search for discovering thousands of groupings of parameter values. Here, we have proposed a fusion approach that can be used to correct parameters in S3VM algorithms.

It can be controlled by a single node in the cluster; it moreover works when multiple processors are required for running the training. For single node running the training, the node loads the training and justification data from the distributed file system of the cluster and studies a model on the training data and intelligences the justification text scraping data. The foremost benefit of the MapReduce model for parameter change is that it can be used for identical big optimizations at small cost.

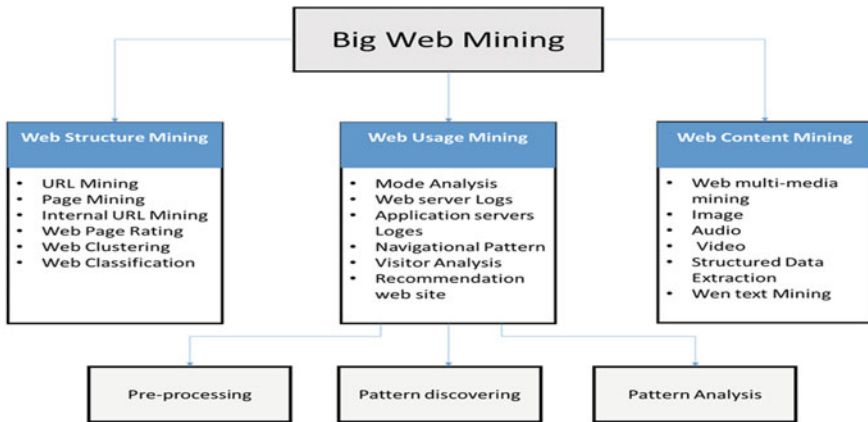


Fig. 1 Big web mining technique

Big web mining technique structures the web document or contents that have extensively disseminated, dynamic and heterogeneous data. This helps creation of multiple domains for information suspicious big text data analysis for prediction terrorism- event classification on the darkweb. Web mining is a data mining technique that helps divide the web mining in different category (as shown in Fig. 1). The different categories of web mining use dissimilar types of properties and behavior for big text data extraction in large amount of data set. It is used for allocating the foremost data, processing approaches and application areas.

The rest of the paper is organized as follows. Section 2 covers the related work. Big text data analysis for prediction terrorism activity using information extraction (IE) and challenges of big darkweb data extraction (WDE) is presented in Sects. 3 and 4, Sect. 5 describes the working of feature selection process, and Sect. 6 presents the result analysis on the finding. Finally, Sect. 7 concludes this paper with future work.

2 Related Work

In this paper, we study the number of CI technique, computational intelligence is the accommodating behavior of distributed, self-organized systems, normal or simulated. The appearance was presented by computational intelligence is that arrangements through customary and artificial systems composed of numerous individuals that coordinate using decentralized mechanism and self-organization. In specific, the discipline efforts on the mutual behaviors that are produced after the local communications of the entities with every additional and through their environment. Instances of schemes studied by computational intelligence are colonies of ants and termites. Selected human objects similarly fall into the domain of computational

intelligence, particularly multi-robot systems, and similarly convinced computer programs that are inscribed to challenge optimization and data investigation difficulties. The training data set is collected from Tumbler social site for assessing the efficacy of methods. Elovici et al. [2] proposed the advance system (advance terrorist detection system) for detecting the terrorist activities is much better than the previous one called terrorist detection system (TDS) and has achieved best score in illicit activity detection. Scanlon [3] proposed the automatically tuned tracking of terrorist recruitment and radicalization activities [4]. Godawatte et al. [5] proposed the identification of ISIS and other terrorist organizations activities working anonymously on darkweb platform. Das Bhattacharjee et al. [6] designed the approach for tracing the techniques of catching the online markets offering illegal services like drug supply, arms trafficking, pornographic contents, human trafficking, organ trafficking. Sachan [7] worked on the identification, tracing and searching of terrorist activities include money laundering, cyptocurrency operations, fundraising, and terrorist-related content sharing [8]. Kaehler et al. [9], Gao and Lopez [10] techniques highlighted the cosmic nature of hidden criminal network for visualizing malicious activities by security agencies.

Bhuyan et al. [11] proposed the taxonomy and comparative structure for cyber scanning and some of the existing detection techniques. Threshold random walk (TRW) presented the network traffic scanning detection model by Jung et al. in [12], presented the well-known algorithm which was designed as part of the Bro [13] IDS. Sridharan et al. in [14] analyzed the efficacy of available techniques and methods such as TRW and Snort and introduced new techniques dubbed as time-based access pattern sequential hypothesis testing (TAPS).

3 Big Text Data Analysis for Prediction Terrorism Activity Using Information Extraction (IE)

Big text data extraction is a technique for classification of a raw data [terrorism activity (usage data)] set in proper manner. In the recent year, number of algorithm work for big text data extraction [15–17]. But they have multiple problems because the data is multidimensional. A number of techniques have been developed for extraction of information for web data set but that algorithm can be used to solve a particular problem. Recent approach can be to solve the data extraction problem but that cannot work for all types of data set. In that case, more number of algorithms will be required for extraction in an intelligent way. In the domain of suspicious big text data analysis for prediction terrorism activity and classification, there are a number of programs that have been classified and extract information from web data set [18, 19].

Big text data analysis for information extraction process performs the post-processing task. The post-processing task is used for all the processing and extraction of information. In the information extraction process, there are a number of packages for classification of the web usage data, email data business enterprise data, criminal

data, etc. These classifications can be as follows: text data, newspaper data, YouTube data, social website data, medical field data, online shopping data, denial of service attack (DoS), distributed DoS attack (DDoS), ransom-related DoS attack (RDoS) attack and so on. The information extraction technique process has been specified in a real-time data that is set for classification during information extraction has been classified the quality of data set that is used for online data extraction [20].

Data information extraction process is simply a way to classify the personalization of data that is used in the near future. The web information extraction process has the capability for evaluation of different types of data set just like denial of service attack (DoS), distributed DoS attack (DDoS) and ransom-related DoS attack (RDoS) that determine the financial fraud [21]. Using the web data, classification and extraction process will predict the user's behavior in various areas like—denial of service attack (DoS), distributed DoS attack (DDoS) and ransom-related DoS attack (RDoS). The information extraction is a broad domain. It will have a number of algorithms that can be used to solve easy and effectively real-time problems. We also apply the machine learning algorithm in the field of information extraction.

We can divide web data extraction approach into two parts:

1. Offline phased data extraction.
2. Online phased data extraction.

This phase contains two foremost components: data preprocessing and knowledge discovering. Initial log preprocessing extracts user's session and identification and will put the essential information into the database. Usually, numerous predefined techniques like data cleaning, session identification and user discrepancy will be essential to complete previous executions for classifications of a web log data using the web mining process.

Knowledge discovered after the data preprocessing step can be incorporated into semantic features of the crops with the mined user session information from the server and user web log. To use RDF model constructed on the JENA framework, RDF is an underpinning for processing metadata. This provides the interoperability among solicitations when conversation machine understands information on the web.

In the online phase of classification, the useful information of the darkweb user reaches the server, the computational intelligence techniques. It associates the semantic knowledge base to distinguish the darkweb user aforementioned transactions.

4 Challenges of Big DarkWeb Data Extraction (WDE)

Big darkweb data extraction is a process of classification of the huge amount of data set, which is called the big data set. This data grows continuously in different formations and number of computational intelligence algorithms that are used for classification of the big darkweb usage data in different manner. Data representation is

also very complicated to understand. Numerous features are essential to be occupied in darkweb usage information classification [22].

In its furthest well-known process, the problem of extracting or classification of information from the darkweb usage mining subsequently threatens its numerous requirements. The major challenge is to find out the information extraction process (Fig. 2).

- The classification process or algorithm is used for improving the accuracy and quality of data set. To ascertain reasonable trade in between building extremely intelligent darkweb data extraction classification technique and the requirement of achieving accurate performance.
- Information extraction process can be classified into huge amount of data called the big data can be classified in less time and efficient manner.
- Using the computational intelligence for information extraction product, the privacy preservation process performs the very effectively consequent proficiency tries to violate user privacy essential to be well timed and suitably diagnosed and counteracted.
- Using the computational intelligence technique performs the training and testing task simply to classify the labeled data set, generated error free, noise-free data for using the behavior predicted by the user.
- Darkweb sources are constantly evolving and structural adjustments occur without a forewarning consequently is unpredictable. Ultimately, in real-global eventualities darkweb information continuously changes and updated data format also

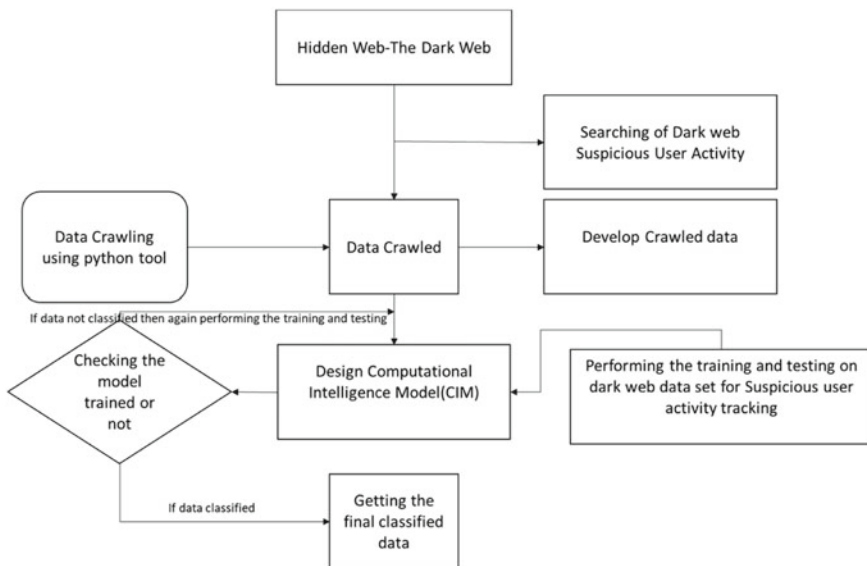


Fig. 2 Flowchart of darkweb data crawling and suspicious user activity determination

changes. No specific algorithm for classifying the particular data set, information source is dynamic.

5 Working of Feature Selection Process

On the basis of selection approach, Feature selection process is extensively characterized into three classes particularly collection, filter and fusion technique. Collection technique calls for a predetermined set of rules to decide the superlative feature subset. Predictive accuracy of the algorithm is used for evaluation. This technique assures higher outcomes, though it is computationally expensive for big data set. For this reason, the collection technique is not frequently preferred. Filter technique selects the feature subset on the premise of intrinsic characters of the facts, self-determining of mining set of rules. It can be carried owing to the availability of statistics with high dimensionality. This recommendation of filter technique is its simplification and high computation performance. Fusion technique combines collector and filter to comprehensive value of each of the techniques. In this method, filter technique is primary instigated to reduce the search space, and then, a collector is simulated to improve the finest feature subset. Our feature selection process works on algorithm necessitates training data for learning purposes. The training data can be whichever, labeled or unlabeled. From the observation of exploiting label information, feature selection algorithms are classified into supervised, unsupervised and semi-supervised algorithms. Supervised feature selection routines labeled data for learning purposes while unsupervised feature selection using unlabeled data.

Algorithm 1: Mapper Function (MF) 1

We applying binary neural network (BNN) based on deep learning concept) performing the training and testing operation using Hadoop tool to implement mapper function

Input: V_1, V_2, W_1, W_2

Output: parsing value O_1, O_2

Step 1: We implement the binary tree

Step 2: Applying the training and testing operation

Step 3: Compute the weight (W_1, W_2)

Step 4: According to training sample (classified) the crawling data (Labeled sample)

Step 5: For performing the testing operation

Algorithm 1: Mapper Function (MF) 2

Step 1: In the second phase, labeled and unlabeled samples determine the error

Step 2: If we get the error, then again apply ordinary least squares regression (OLSR) and complete the testing process

Step 3: Control all operation

- Step 4: Performing the testing operation
- Step 5: Checked the sample stage if not present unlabeled sample
- Step 6: Stop condition

Algorithm 2: Infiltrate Function

Applying backpropagation with learning vector quantization (LVQ). We apply infiltrate function on middle operational data set

Input: O_1, O_2, V_1, V_n

Output: O_2, O_2

- Step 1: Applying the LVQ approach on testing sample
- Step 2: For (Select appropriate range)
- Step 3: Compute the minimum distance (Using Euclidean distance)
- Step 5: For (wait for process completion)
- Step 6: Terminated the condition
- Step 7: Return O_2, O_2 as a pairs

Algorithm 3: Reducer Function

We combine all output generated infiltrate

Input: O_2, V_2

O_2 : This is key generated by the Infiltrate function

V_2 : This is value generated by the Infiltrate function

Output: O_2, V_2

- Step 1: Preforming the sorting operation (on the key value pair)
- Step 2: If (check the correct pair)
- Step 3: Determine the proper classified data (using testing algorithm)
- Step 4: Return the correct output
- Step 6: Terminated the condition
- Return O_2, V_2 (Fig. 3).

Darkweb classification is used for the online modeling and offline modeling constructing the user profile model since relations are the offline quantity of modeling. For novel users and different activity intensification, the present users' attention modification over time and consequently user's determination prepare novel transactions with the darkweb application, to use a method to increase data classification by changing model complete different transactions. Through this technique deliberate, the behavior of user meanwhile creates the offline model and computes the probability of passing on a user to additional constellations. At the major, stimulating of every constellation nearby every class in a predefined interval is designed by counting the amount of request sent to every class from all constellations. Then, to compute the probability of interest of every user about every grouping by including the amount of demand subsequently, a user is concerning the probability of constellations interesting and lastly to allocate a user to a constellation that has concentrated probability. The phase of attributing groups to constellations can be complete by

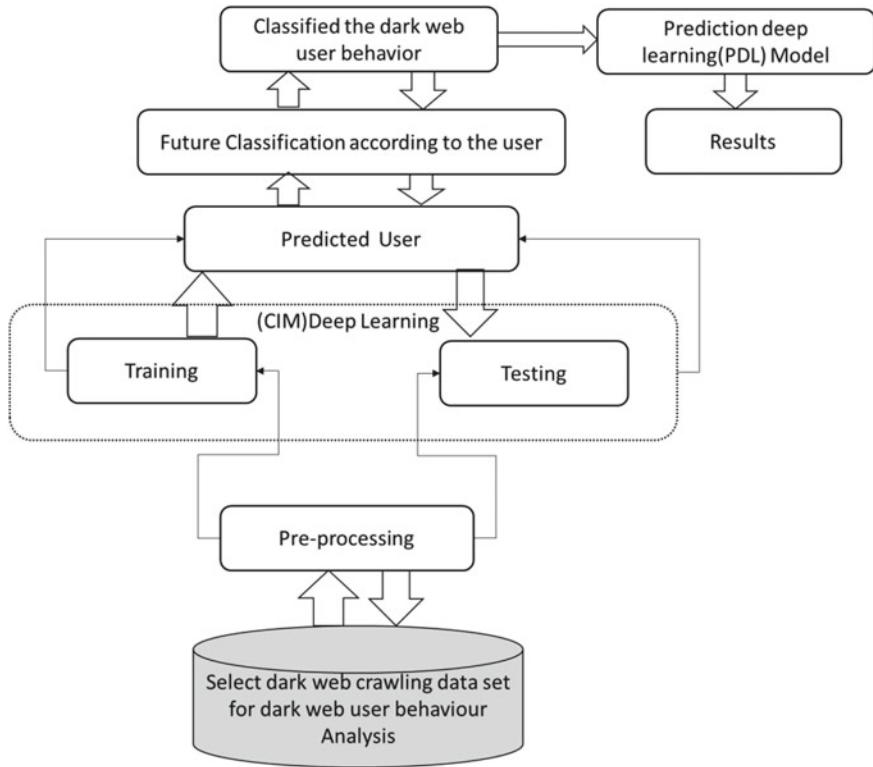


Fig. 3 Suspicious big text data analysis for prediction using computational intelligence model

since the initial constellations accomplished to complete the Deep learning and by the number of transactions completed by the users of every constellation through a group domain subsequently producing the static model in offline mode.

Every group is connected to the constellations through a probability value contemporaries the percentage of the number of requirements sent to the groupings users of the patterns overall constellations. Though this type of construction, association is not entirely precise, and then, if the duration of time be extended and the number of transactions is huge sufficient, this can be consistent. Additionally, there are particular methods for discovery constellations; this is deliberate the satisfied types of data. The frequent item set is an instance of discovery relation amongst user constellations and content groups.

For computer relative among constellations and group G_l in ontology domain, D_l is definite as the creation of showed users and C_i as initial part created by a deep learning algorithm. Specified the complete amount of users fitting to every part subsequently the part O_l , the complete quantity of requirements of D_q in constellation I_l for every category the requirements of constellation I_l for every group is the entire amount of requirements since the constellation separated by an entire number of users inside

the cluster. It is standardized necessities sent for each group over complete quantity of users privileged each constellation.

$$HI_{lm} = \sum_{D_q \in I_l}^n HD_{qm} / O_l \quad (1)$$

A matrix of constellation group signifies connection amongst constellation and groups. Every cell characterizes probability PI_{lm} that is comparative among group P_m , and constellation I_l is considered:

$$PI_{lm} = HI_{lm} / \sum_{l=0}^n HI_{lm} \int O_l \quad (2)$$

Before, to discover the alteration of users' benefits by since transactions completed subsequently constellation. If the requirements that user has directed complete transactions for satisfied classes are dissimilar through the classes that his constellation fits, the constellation of user determination modification. In the additional expression, the probability that a user goes to a constellation multiplied by including the number of times every user directs demand for each group after building the offline model.

$$PDI_{ql} = \sum_{D_q \in I_l} HI_{lm} * PI_{lm} / O_l \int O_l \quad (3)$$

Process transfer each user to the constellation that user has determined probability that is fit to that. Consequently, if the common of a user's demand is aim at the class that has not a high association through her constellation, user determination drops in an additional constellation. It deliberates the deviations in the performance of users subsequently offline exhibiting and not essential to put on deep learning algorithm on complete data. Consequently, the model used for active modeling deprived of scalability deficiencies in customary online models.

Classification of Semi-supervise Learning

We analyze specific of the modern process for semi-supervised learning. While cannot designate completely accessible approaches, instead of emphasis on three classes of process in graph-based process, modification of illustration process and margin created regularization-based process. Each of these process (moreover implicitly or explicitly) are different belongings to a regularization process of semi-supervised learning. In supervised learning, a regularization process indications one to optimize the observed risk (training error) plus a further term that limits the complexity of the decision rule:

$$\min_{\alpha} \sum_{i=1}^L l(f(x_i, \alpha) y_i) + \gamma \Omega(\alpha) \tag{4}$$

where $l(\cdot)$ is a loss meaning training the consequence for erroneously labeling a training instance, and α encodes the parameters of the resolution function. In semi-supervised learning, unique technique is to do precisely the similar as in supervised learning, but enhance one further term that encodes the hypothesis one has complete about the data using unlabeled samples. One might wish to encode the cluster hypothesis. This leads one to minimize:

$$\min_{\alpha} \sum_{i=1}^L l(f(x_i, \alpha) y_i) + \gamma \Omega(\alpha) + \lambda \tilde{A}(x^*, \alpha) \tag{5}$$

$\tilde{A}(x^*, \alpha)$ regularization point wise

$$\tilde{A}(x^*, \alpha) = \sum_{i=1}^U l^*(f(x_i^*), \alpha) \tag{6}$$

where $l(\cdot)$ is a function that encodes the assumption of excellent. To currently deliberate particular precise semi-supervised learning algorithms, semi-supervised learning approaches can be separated in classes of algorithms, considering the used approach.

Combinative Optimization on Darkweb User Activity Using Computational Intelligence Techniques

Unsupervised support vector machine created on completely pointed sets to make it learn following freedom, since, in actual-world problems for pointing the data can be spendthrift as well as time captivating. Following are the given details, plane data is supposed to be reserved again and again but in case of unobserved learning, it is not accurate always. Hence, in today’s time, classifiers have adopted a new policy of learning by manipulating the partially training cases collaborating with well pointed and plane data as well. So here we deliberate the semi-supervised support vector machine cases association labeled and unlabeled data in the proposed of learning classifiers.

Define a parameter

\bar{T} = training set now, consistent to the training set \bar{T} well defined by

\tilde{T} = specified a data set

x_p = the labeled input

y_p^+ = positive class with probability

y_p^- = negative class with probability

\tilde{x}_p = the label of input is unidentified

\tilde{x}_j = unknown label of input
 $i = 1, \dots, p$
 $J = 1, \dots, q$, an added doublet parameter.

Supposing we are a particular data set \bar{T} having P well pointed data (the making to learn set) along with freedom and Q plane data (the examining set)

$$\bar{T} = \{(x_p y_p^+, y_p^-) \dots (x_p y_p^+, y_p^-)\} \cup \{\tilde{x}_p, \dots, \tilde{x}_q\} \tag{7}$$

y_i^+, y_i^- Greeter than equal to zero and $y_i^+ + y_i^-$ less than equal to one.

The ambiguity is similarly labeled precisely by probability. Additional precisely, for each coupling the data set \tilde{T} having instruction giving set \bar{T} , the most initiatively made change is to have plane data in operating pair mood in \tilde{T} .

\tilde{x}_j = Unknown label of input where y_j^+, y_j^- and $y_j^+ + y_j^-$.

In the similar indication, to articulate the S^3VM along possibility for making the continuous plane.

$$f(x) = \langle w.x \rangle + b \tag{8}$$

As follows:

Increasing the boundary is same however to squeeze $\|w\|/2$ in specific smartness, that is very vital duration of the transparent working.

6 Results Analysis

Number of steps to perform pre-processing data set:

1. Here, with this approach, the redundant data is removed, and data is stored in a proper mode.
2. To format data sets, its arrangement is in a proper way is highly required.
3. If the information is not according to the user knowledge, then the whole process will be repeated.
4. Now, it is needed to remove the unwanted information.
5. Here, we can identify the missing values and note them to remember them for further usage.
6. There are some transactions performed by the user that are ambiguous in nature so we will try to determine them.

Steps for working of algorithms are as follows:

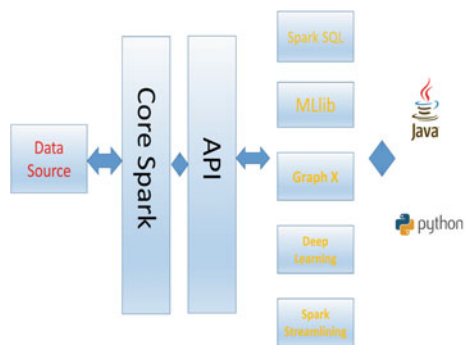
- Firstly, we have to choose a training sample and then perform the training task using parameters and related kernel function. The algorithm used here is S3VM.
- Now, we are selecting both the data sets that are labeled and unlabeled. Here, we are using active learning technique.
- We combine all the samples of data sets to generate a new data sets for training sets.
- We are having the training sets through which we will generate new classifiers.
- If there are unlabeled sample in label sample, then we will classify them again.
- If we are not getting the desired results in this process, then we have to do repeat all the processes.
- If we are getting the desired results, then stop training process and exit.

Identification of the Non-determine sample

- HADOOP is a common data analytics framework based on distributed execution of cluster. It executes for increased speed and data progression completed by MapReduce. HADOOP cluster runs on the top and access HADOOP distributed file systems and can progress structured data in hive and streaming data from, Kafka, flume, HDFS, Twitter dissimilar executions as well.
- The collected complete transitional values are connected with the parallel middle key, and MapReduce library groups are permitted. User can write the reducer function according to our requirement and a set of values for that key to use the lesser value for integration. Reducer can produce the zero and one output. The middle value helps the reducer in completing the iteration. MapReduce measurement object comprises of labels of input–output files and possible alteration of value. Our proposed approach can be used to develop the intelligent map reducing model for classification of the information at proper level (Figs. 4 and 5; Tables 1 and 2).

Complexity Evaluation: After simulation of our proposed algorithm system, the performance of the classification in terms of accuracy, error rate, space complexity

Fig. 4 HADOOP spark



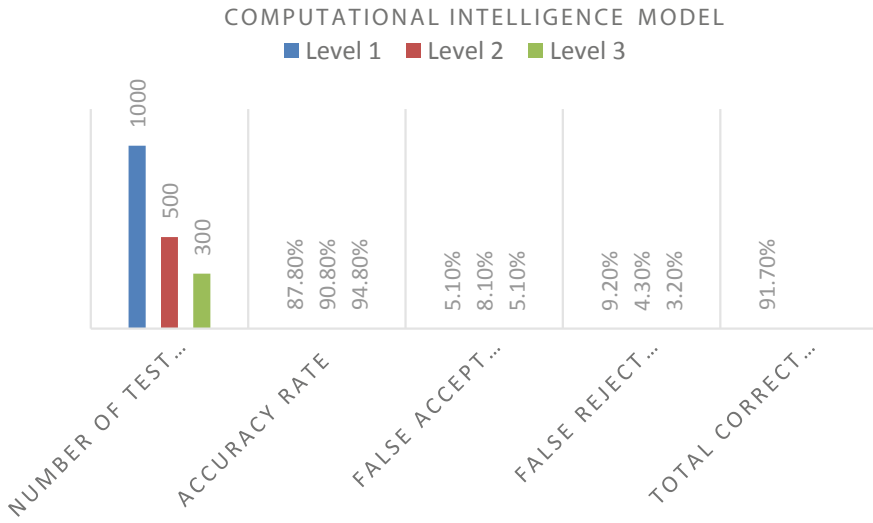


Fig. 5 Result evaluation

Table 1 Dark web data set using CIM

Cyber terrorist attack (Darkweb data set)	Parameter	Computational intelligence algorithm (attacking data classification accuracy in %)	
		Existing algorithms	Proposed algorithm
Crawled data set 1	FreeBay, Havana/Absolem, Haven, Horizon, Hydra, Ironclad, Kiss, cannabis, cocaine, darknet, drug markets, opioids, stimulants	91.80	92.33
Crawled data set 2	1776, Abraxas, Agape, Agora, Alpaca, cocaine	85.00	92.83
Crawled data set 3	Middle Earth, Mr Nice guy 2, Nucleus, Onionshop, Outlaw Market, Oxygen, Panacea	86.97	88.58
Crawled data set 4	Sheep, Silk Road 1, Silk Road 2, Silk Road Reloaded (I2P), Silkstreet	83.93	91.53
Crawled data set 5	Doge Road, Dream Market, Drugslist, East India Company, Evolution	92.91	94.61

Table 2 Experiment results

Parameter	Level 1	Level 2	Level 3
Number of test samples	1000	500	300
Accuracy rate (%)	87.8	90.8	94.8
False accept rate (FAR) (%)	5.1	8.1	5.1
False reject rate (FRR) (%)	9.2	4.3	3.2
Total correct classification rate (%)	91.7		

Table 3 Performance summary

S. No.	Parameters	SVM	SVMG-RBF	TSVM	S3VM	Fusion S3VM
1	Accuracy	High	Low	High	High	High
2	Error rate	Low	High	Low	Low	Very low
3	Memory consumption	Low	High	High	High	High
4	Time consumption	High	Low	Low	Low	Very low

and time complexity is assessed and related with SVM, SVMG-RBF, TSVM and S3VM. The comparative performances of all the algorithms are summarized in Table 3.

- Agreeing to the comparative performance, proposed algorithm performed proficiently and accurately due to high training and testing rate and improving the error rate throughout the prediction. Furthermore, the space complexity of the proposed is less, however, the performance of the system leaked in positions of time complexity. That requires more time for accurate learning.
- Time Complexity: The proposed model is assessed in relation to the time complexity. In addition, we analyzed computation time of the classification model through the current classification deprived of MapReduce methods and proposed a technique with mapper reduce proposed based classification. It is separate, as the time and accuracy for the classification process in comprehensive data with the present methods is associated with the mapper reduce-based grouping technique.

7 Conclusions and Future Work

Using our proposed model improving the information extraction process and find out the user unauthorized activity according to search behavior, for improving the prediction accuracy. Our model is also using neural network and AI concept to improve the search perfection to apply the S3VM with the suing of deep learning algorithm using probability distribution technique for increasing the search behavior accuracy. Proposed model is improved and modified the search space, which has

increased the domain knowledge search behavior. Read the user preference, create the model for analysis of the click stream, browsing history. We find out the numeric data performs the machine learning technique that create the model. Suspicious big text data analysis for prediction—on darkweb user activity using computational intelligence techniques. We try to implement in real-time environment in the near future.


References

1. O. Chapelle, V. Sindhwani, S.S. Keerthi, Optimization techniques for semi-supervised support vector machines. *J. Mach. Learn. Res.* **9**, 203–233 (2008). <https://doi.org/10.1007/s11277-020-07143-2>
2. Y. Elovici et al., Content-based detection of terrorists browsing the web using an advanced terror detection system (ATDS), in *Terrorism Informatics. Integrated Series in Information Systems*, vol. 18, ed. by H. Chen, E. Reid, J. Sinai, A. Silke, B. Ganor (Springer, Boston, 2008). https://doi.org/10.1007/978-0-387-71613-8_17
3. J.R. Scanlon, M.S. Gerber, Automatic detection of cyber-recruitment by violent extremists. *Secur. Inform.* **3**, 5 (2014). <https://doi.org/10.1186/s13388-014-0005-5>
4. R.M. Silva, I. Zelinka, An overview of cyber insecurity and malicious uses of cyberspace, in *AETA 2017—Recent Advances in Electrical Engineering and Related Sciences: Theory and Application. AETA 2017. Lecture Notes in Electrical Engineering*, vol. 465 (Springer, Cham, 2018). https://doi.org/10.1007/978-3-319-69814-4_2
5. K. Godawatte, M. Raza, M. Murtaza, A. Saeed, Dark web along with the dark web marketing and surveillance, in *2019 20th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)*, Gold Coast, Australia (2019), pp. 483–485. <https://doi.org/10.1109/PDCAT46702.2019.00095>
6. S. Das Bhattacharjee, B.V. Balantrapu, W. Tolone, A. Talukder, Identifying extremism in social media with multi-view context-aware subset optimization, in *2017 IEEE International Conference on Big Data (Big Data)*, Boston, MA (2017), pp. 3638–3647. <https://doi.org/10.1109/BigData.2017.8258358>
7. A. Sachan, Countering terrorism through dark web analysis, in *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)*, Coimbatore (2012), pp. 1–5. <https://doi.org/10.1109/ICCCNT.2012.6396055>
8. G. Bovenzi, G. Aceto, D. Ciunzo, V. Persico, A. Pescape, A big data-enabled hierarchical framework for traffic classification. *IEEE Trans. Netw. Sci. Eng.* <https://doi.org/10.1109/TNSE.2020.3009832>
9. R. Kaehler, O. Hahn, T. Abel, A novel approach to visualizing dark matter simulations. *IEEE Trans. Vis. Comput. Graph.* **18**(12), 2078–2087 (2012). <https://doi.org/10.1109/TVCG.2012.187>
10. B.J. Gao, J.A. Lopez, LIGHT: enabling instant communication for web surfers with momentary needs, in *2018 IEEE International Conference on Big Data (Big Data)*, Seattle, WA, USA (2018), pp. 5322–5324. <https://doi.org/10.1109/BigData.2018.8622577>
11. M.H. Bhuyan, D. Bhattacharyya, J.K. Kalita, Surveying port scans and their detection methodologies. *Comput. J.* **54**(10), 1565–1581 (2011)
12. J. Jung, V. Paxson, A.W. Berger, H. Balakrishnan, Fast portscan detection using sequential hypothesis testing, in *2004 IEEE Symposium on Security and Privacy (IEEE, 2004)*, pp. 211–225
13. I. Bro, Homepage: <https://www.bro-ids.org> (2017)
14. A. Sridharan, T. Ye, S. Bhattacharyya, Connectionless port scan detection on the backbone, in *25th IEEE International Performance, Computing, and Communications Conference, IPCCC 2006 (IEEE, 2006)*, p. 10

15. K. Barhanpurkar, A.S. Rajawat, P. Bedi, O. Mohammed, Detection of sleep apnea & cancer mutual symptoms using deep learning techniques, in *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Palladam, India (2020), pp. 821–828. <https://doi.org/10.1109/I-SMAC49090.2020.9243488>
16. A. Singh Rajawat, S. Jain, Fusion deep learning based on back propagation neural network for personalization, in *2nd International Conference on Data, Engineering and Applications (IDEA)*, Bhopal, India (2020), pp. 1–7. <https://doi.org/10.1109/IDEA49133.2020.9170693>
17. A.S. Rajawat, O. Mohammed, P. Bedi, FDLM: fusion deep learning model for classifying obstructive sleep apnea and type 2 diabetes, in *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Palladam, India (2020), pp. 835–839. <https://doi.org/10.1109/I-SMAC49090.2020.9243553>
18. A.S. Rajawat, A.R. Upadhyay, Web personalization model using modified S3VM algorithm for developing recommendation process, in *2nd International Conference on Data, Engineering and Applications (IDEA)*, Bhopal, India (2020), pp. 1–6. <https://doi.org/10.1109/IDEA49133.2020.9170701>
19. A.S. Rajawat, P. Upadhyay, A. Upadhyay, Novel deep learning model for uncertainty prediction in mobile computing, in *Intelligent Systems and Applications. IntelliSys 2020. Advances in Intelligent Systems and Computing*, vol. 1250 (Springer, Cham, 2021). https://doi.org/10.1007/978-3-030-55180-3_49
20. S. Mandal, S. Biswas, V.E. Balas, R.N. Shaw, A. Ghosh, Motion prediction for autonomous vehicles from lyft dataset using deep learning, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India (2020), pp. 768–773. <https://doi.org/10.1109/ICCCA49541.2020.9250790>
21. S. Mandal, V.E. Balas, R.N. Shaw, A. Ghosh, Prediction analysis of idiopathic pulmonary fibrosis progression from OSIC dataset, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India (2020), pp. 861–865. <https://doi.org/10.1109/GUCON48875.2020.9231239>
22. M. Kumar, V.M. Shenbagaraman, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering*, ed. by M.N. Favorskaya, S. Mekhilef, R.K. Pandey, N. Singh (Springer, Berlin, 2020), pp. 765–773. ISBN 978-981-15-4691-4

Development of Prediction Models to Determine Compressive Strength and Workability of Sustainable Concrete with ANN



Priyanka Singh , Saurav Bhardwaj, Saurav Dixit, Rabindra Nath Shaw, and Ankush Ghosh

Abstract Concrete is the widely used composite material to build infrastructures that should remain durable and serviceable. This study focuses on sustainable concrete produced by replacing cement with waste products (fly ash and slag) to obtain good mechanical long-term and desirable durability. The use of fly ash and slag in concrete in place of cement lowers global energy demand and saves money on the verge of depletion. Fly ash provides increased workability, mechanical and durability properties and a wide range of advantages, including decreased strain on natural resources and a lower CO₂ footprint. Fly ash in concrete structures is used as a partial substitute for Ordinary Portland Cement. Experimental work about concrete contributes to a waste of resources, time and money. Nominal concrete is a heterogeneous mixture of cement, water, coarse and fine aggregates. Therefore, for efficient use of such materials in various engineered structures, a clear understanding of such complex behavior is important. In this paper, the ANN is used to execute and provide output for compressive strength and workability of concrete, which we are predicting by use of 103 datasets obtained from available technical literatures. The focus is to find out the ideal equation for workability and compressive strength for concrete with the help of ANN.

Keywords Artificial neural network (ANN) · Green concrete · Sustainable concrete · Compressive strength · Workability · Regression model

P. Singh (✉) · S. Bhardwaj
Department of Civil Engineering, Amity School of Engineering & Technology, Amity University
Uttar Pradesh, Noida, Uttar Pradesh, India

S. Dixit
Pioneer Research and Consultancy Service, Gautam Buddha Nagar, Uttar Pradesh, India

R. N. Shaw
School of Electrical, Communication & Electronics Engineering, Galgotias University, Greater
Noida, Uttar Pradesh, India

A. Ghosh
School of Engineering and Applied Sciences, The Neotia University, Sarisha, West Bengal, India

1 Introduction

Concrete is one of the world's widely utilized engineering construction materials. Its distinctly compressive strength is generally recognized as the primary indicator of concrete quality performance. The blending of cement, water, fine aggregates and coarse aggregates is used to attain concrete strength. Sustainable concrete is the composite material composed of cement and other kinds of binding agent, which is used to improve the mechanical characteristics along with fine aggregates (FA), coarse aggregates (CA) and admixtures like fly ash, silica fume, rice husk ash, marble dust, metakaolin, etc. [1–3] Concrete blends of various combinations must be formulated by using different proportions.

Recent developments in environmental waste treatment regulations such as fly ash and slag have caused concerns about using waste as building materials to replace Portland concrete cement partially. The substitution percentage varies from roughly 20% (low volume FA) of the overall cement materials to more than 50% (high volume FA) [4–6]. FA and slag were widely used for cement replacement with concrete. Furthermore, FA may be substituted by up to 60% if the early intensity is not a necessary element. It is well established that fly ashes were usually counterproductive to concrete capacity.

Ash is precipitated by the exhaust pollution of coal-fired power stations electrostatically or mechanically and used in the early stages of mass concrete to reduce hydration. FA may have more power and longevity when it is made with lower water-to-binding ratios. There is a nonlinear correlation between ingredients of concrete and its compressive resulting strength. In this situation, traditional compressive strength assessment strategies require the development of several concrete samples for each concrete grade; then, the samples are cured for 28 days under normal circumstances [7–9]. Using non-destructive techniques to calculate and identify concrete compressive strength can contribute to tremendous savings in time, material costs and labor. The quest for a simple, effective and non-destructive solution offers an alternative to laborious destructive tests. It provides an efficient method that is non-destructive and forecasts the compressive strength of concrete as an alternative to laborious destructive testing. Concrete should be workable well as it should possess desirable compressive strength. The workability and compressive strength of concrete are inversely proportional to each other. If the workability of concrete is more (water content is high), it will drastically reduce its compressive strength (CS). Hence, in this research, predictive models have been developed to optimize concrete compressive strength with good workability through the slump and flow value of concrete.

Nonlinear regression is not used for prediction because they determine the bond b/w independent and dependent variable by fitting a linear equation. Multiple regression analysis is a process to analyze the dependent variable by the associated dependent variable.

A usual equation for multiple linear regression is

$$N = A_0 + A_1Y_1 + A_2Y_2 + \dots$$

when there are some multiple nonlinear dependencies, the logarithmic function is used, where A_0, A_1, A_2, \dots are the coefficients, and Y_1, Y_2, Y_3, \dots are the independent variables.

In our hypothesis, these aspects have a nonlinear association with the compressive strength of concrete.

2 Artificial Neural Network

To imitate the human brain, artificial neural networks (ANN) have been created. In contrast to the human brain, an ANN that is basic and limited in size has certain powerful characteristics in the transmission of awareness and information due to its resemblance to the human brain. Therefore, for engineering purposes, ANN may be an effective method. It is assumed that the first work on ANN began in 1943. Then, as a second strike, a computer named the perceptron was designed in 1958, and it worked just like the human mind [10, 11].

An artificial neuron comprises five main components: inputs, weights, sums, activation functions and outputs. The biological nervous system inspires these components. The links between components primarily define the network structure. By changing the correlations (weights) between components, neural networks can be trained to perform a specific role.

Usually, neural networks are trained to generate a given goal output. Neural networks were qualified to carry out complex functions that are difficult for traditional computers or humans in different areas [12, 13]. ANN has been trained in recent years to overcome challenges challenging for traditional machines or human brains. During the past three decades, ANN implementations have been used in many different fields of research successfully. We may build ANN models with innumerable combinations with related parameters [14].

The basic methodology for developing models on ANN is to train and test the datasets. The pattern of material behavior is applied in the eligible ANN structures classified as a material model. In addition to being able to replicate the experimental results, the trained ANN structures would also be able to use their generalization ability to approximate the results in other experiments [15, 16]. The purpose of this analysis is to propose ANN models for the prediction of concrete compressive strength with its workability.

The concrete with admixtures has a remarkable tolerance to segregation and deformation. The combination of concrete compositions should be so that the composition does not segregate or bleed [17, 18]. The remarkable segregation resistance and deformation property are found in nominal concrete. The proper mixing of concrete provides preventative steps against bleeding and segregation as it fills the voids to a densely compact structure. So, to predict the behavior of concrete, ANN models are prepared. The ANN models are assisted by computational elements connected to predicting and optimizing their constituents' strength and proportion. Fundamentally, ANN components are like the brain neurons, which comprise various computing

components in layers. ANN operations, as seen in construction technology to assess drying shrinkage of concrete during transportation of ready-mixed concrete, are seen as recent advances. The workability of concrete can be increased with the addition of the materials like fly ash, slag, metakaolin rice husk ash and many other materials. Further, it increases the CS of concrete in the long run. The mixture of concrete designs created from the artificial neural network ought to be optimal with cement and water, leading to higher strength and a relatively better economy and ecology [19, 20].

3 Experimental Setup

The ANN model is an efficient methodology that offers feasible solutions to problems that traditional strategies, such as multiple regression models, do not address and may not invalidate current techniques. Preliminary models were created by the data of several mixed design proportions of concrete based on the experimental observations. 30% of data was arbitrarily used for training out of the whole data to predict the compressive strength of concrete and its workability. In the development of models through ANN, nine inputs and one output were finalized. The input parameters constituted cement, coarse aggregate, fine aggregate, fly ash, slag, plasticizers, slump value, water content and flow value. The prediction obtained by the generated model provides 28 days compressive strength of concrete and its workability. The aim of this study is to explore the artificial neural networks with large volumes of experimental literary data and to predict the compressive strength and workability of concrete with fly ash and slag.

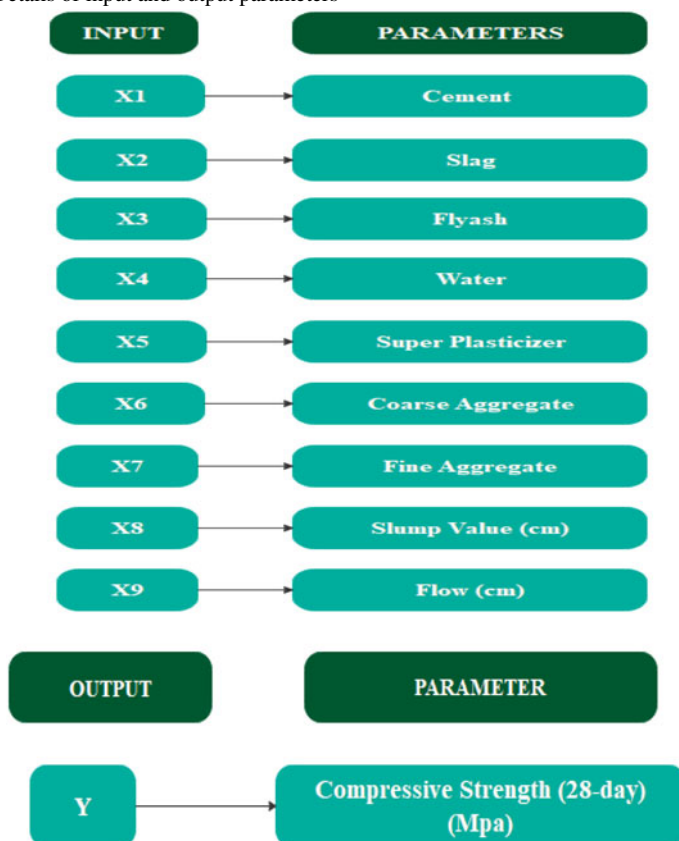
4 Artificial Neural Network Model Design

A comprehensive model is framed to predict the 28 days compressive strength of concrete and its workability, dependent on the different proportions of design mixes. One set of data (30%) is taken for training and the remaining for testing. Artificial neural network (ANN) is implemented in two steps: training and testing. During training time, the input parameters are put in an ANN input layer, while the corresponding value is given as a target for desired output at the output layer.

We used the dataset of concrete to determine the regression model through commercial software named MINITAB (Tables 1 and 2).

Table 1 Composition of concrete

Denotations	Compound	Min.	Max.
A	Cement (kg/m ³)	137	149
B	Slag (kg/m ³)	0	193
C	Fly ash (kg/m ³)	0	260
D	Water (kg/m ³)	166	240
E	Super plasticizer (kg/m ³)	0	19
F	Coarse aggregate (kg/m ³)	735	1049.9
G	Fine aggregate (kg/m ³)	640.6	902
H	Slump (in cm)	0	29
J	Flow (in cm)	20	78

Table 2 Details of input and output parameters

5 Result and Discussion

The prediction value of compressive strength and workability of concrete was determined with ANN. The value that we are predicting is termed as dependent value, and the input parameter to find the prediction of dependent value is termed as an independent value. In this model, the dependent variable is workability and compressive strength.

First, we are checking all the fitted line plotted graphs for each compound to compression strength. A normal probability plot of residuals is plotted to validate normally distributed assumptions. Residuals versus fits plot verifies that residuals have a constant variance. The residual histogram is used to assess data skew and outlier presence in the data. Residual versus data order verifies presumption of uncorrelated residuals. The compressive strength values expected after model analysis are checked for accuracy to assess errors (Figs. 1, 2, 3, 4, 5, 6 and 7; Table 3).

These are a few coefficients value obtained from dataset. It is a measure of specified characteristics for a component used to obtain ideal concrete.

In this Table 4, the model summary has been shown, in which the value of prediction is according to the *R*-squared method. *R*-squared methods are statically measured in which they state how many values are closed to the fitted regression line. The *R*-squared prediction value coming out is 88.45%, which seems a good prediction value (Figs. 8 and 9).

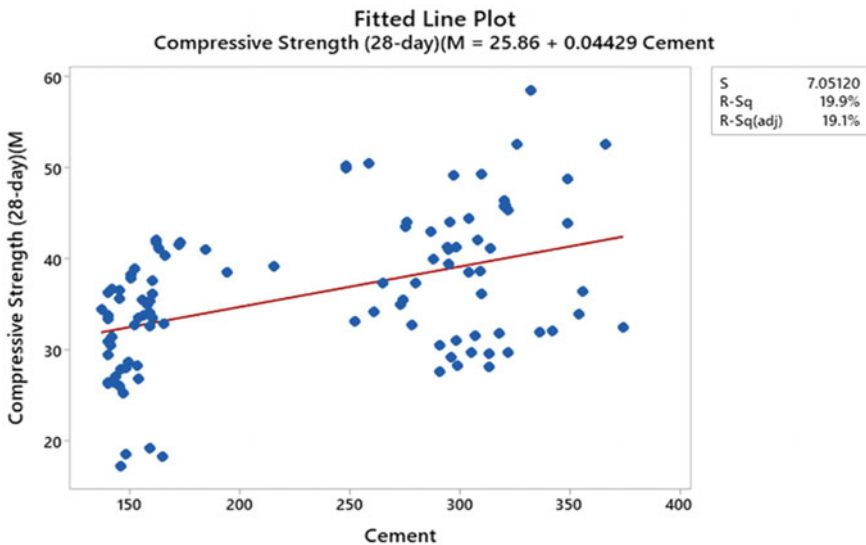


Fig. 1 Fitted line plot between compressive strength (MPa) versus cement content

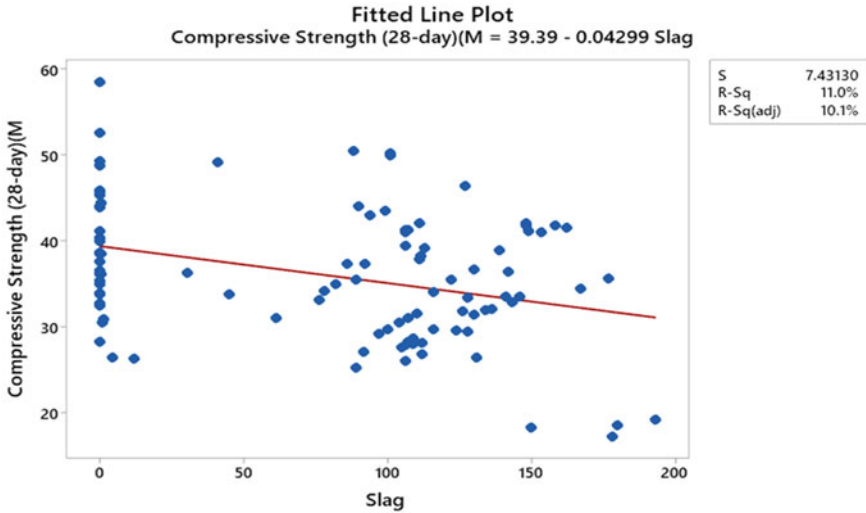


Fig. 2 Fitted line plot between compressive strength (MPa) versus slag

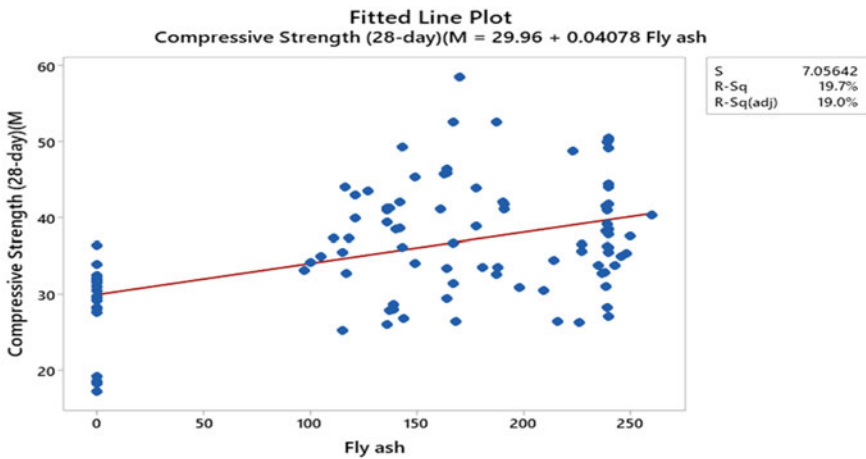


Fig. 3 Fitted line plot between compressive strength (MPa) versus fly ash

The value was plotted in a graph according to the instances given to forecasting the response value. The residual and percentage of the normal probability plot are on the fitted line in the residual graph. The line has been drawn according to the given instance, which is interpreted through software used. Pareto chart is also giving the value that mainly affects the compressive strength through the response of every instance used to prediction, that is, which ingredient is very prominent for compressive strength and which compound presence does not affect the compressive strength of concrete.

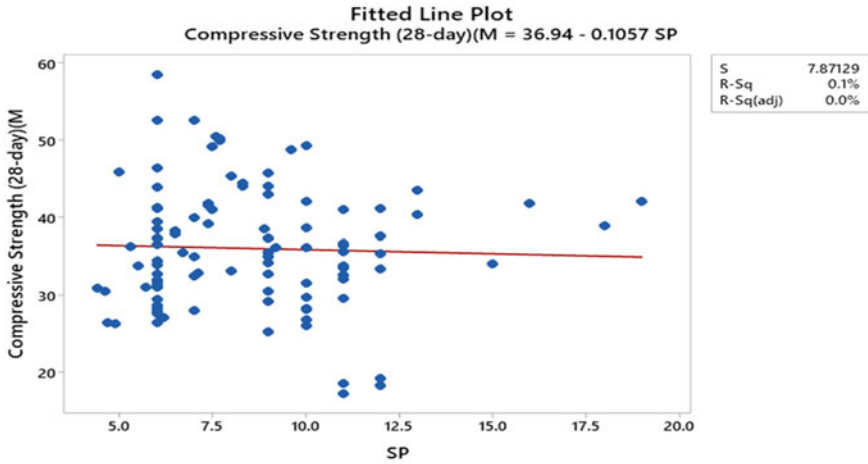


Fig. 4 Fitted line plot between compressive strength (MPa) versus superplasticizer

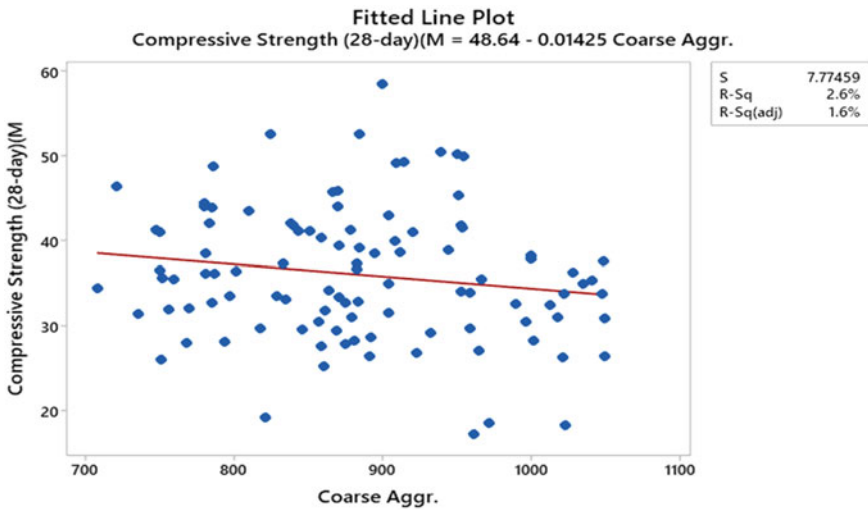


Fig. 5 Fitted line plot between compressive strength (MPa) versus coarse aggregate

Regression equation:

$$\begin{aligned} \text{Compressive strength(28 days)} &= 140.3 + 0.0593 \text{ Cement} - 0.0323 \text{ Slag} \\ &\quad + 0.0468 \text{ Fly ash} - 0.02336 \text{ Water} + 0.034 \text{ SP} \\ &\quad - 0.0548 \text{ Coarse Aggregate} - 0.0379 \text{ Fine aggregate} \\ &\quad - 0.2361 \text{ Slump(cm)} + 0.0848 \text{ Flow(cm)} \end{aligned}$$

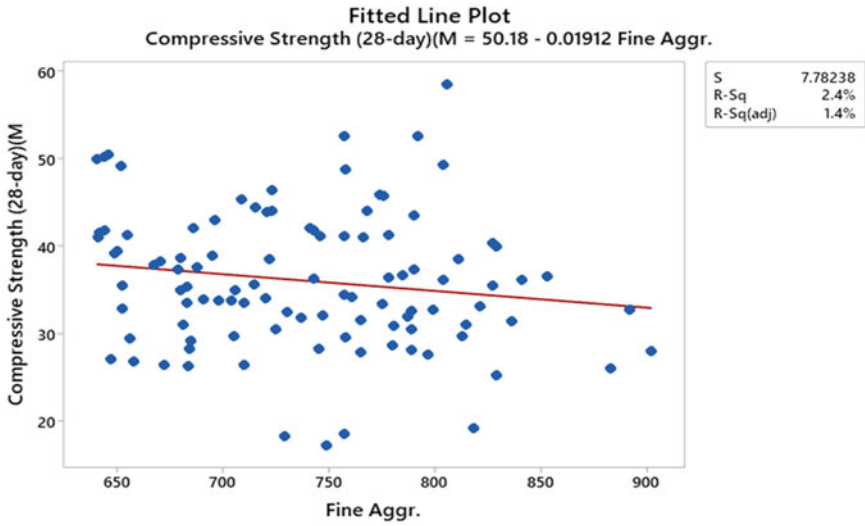


Fig. 6 Fitted line plot between compressive strength (MPa) versus fine aggregate

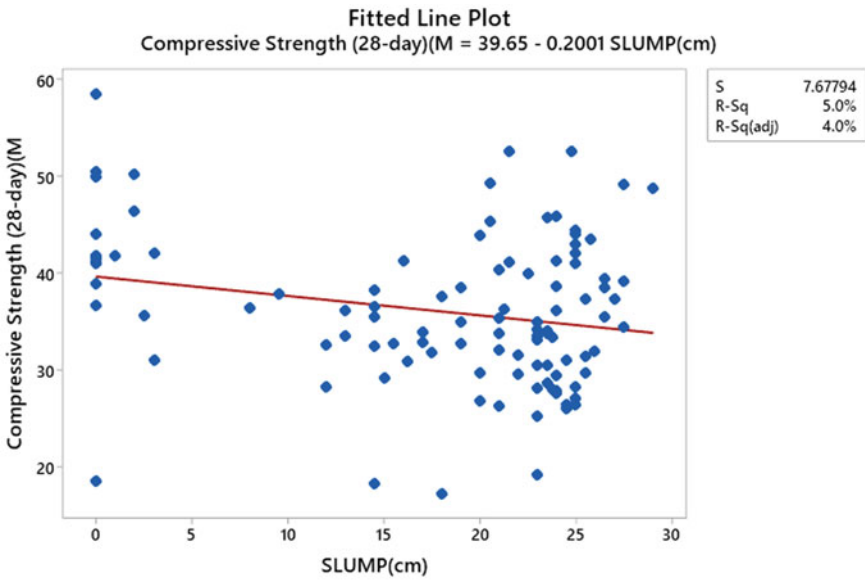


Fig. 7 Fitted line plot between compressive strength (MPa) versus slump

Table 3 Coefficient for every component

Coefficients					
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	140.3	67.5	2.08	0.040	
Cement	0.0593	0.0216	2.74	0.007	48.88
Slag	-0.0323	0.0300	-1.08	0.285	55.31
Fly ash	0.0468	0.0220	2.13	0.036	59.29
Water	-0.2336	0.0698	-3.34	0.001	33.44
SP	0.034	0.130	0.27	0.792	2.23
Coarse aggregate	-0.0548	0.0260	-2.11	0.037	88.50
Fine aggregate	-0.0379	0.0273	-1.39	0.168	50.23
Slump (cm)	-0.2361	0.0759	-3.11	0.002	7.41
Flow (cm)	0.0848	0.0441	1.92	0.058	10.07

Table 4 Model summary and analysis of variance

Model summary					
S	R-sq	R-sq (adj)	R-sq (pred)		
2.46533	90.98%	90.11%	88.45%		
Analysis of variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	9	5701.42	633.492	104.23	0.000
Cement	1	45.62	45.624	7.51	0.007
Slag	1	7.03	7.027	1.16	0.285
Fly ash	1	27.49	27.492	4.52	0.036
Water	1	67.98	67.978	11.18	0.001
SP	1	0.43	0.427	0.07	0.792
Coarse aggregate.	1	27.09	27.087	4.46	0.037
Fine aggregate	1	11.73	11.730	1.93	0.168
Slump (cm)	1	58.75	58.747	9.67	0.002
Flow (cm)	1	22.46	22.463	3.70	0.058
Error	93	565.24	6.078		
Total	102	6266.66			

After using the available data, we have got the prediction value for compressive strength and got a regression equation to produce the ideal composition of concrete to achieve great strength. Simultaneously, regression analysis for slump and flow value was performed to determine the ideal equation to make concrete compact and durable. Regression analysis for slump value versus all other parameters was performed with same dataset (Tables 5 and 6; Figs. 10 and 11).

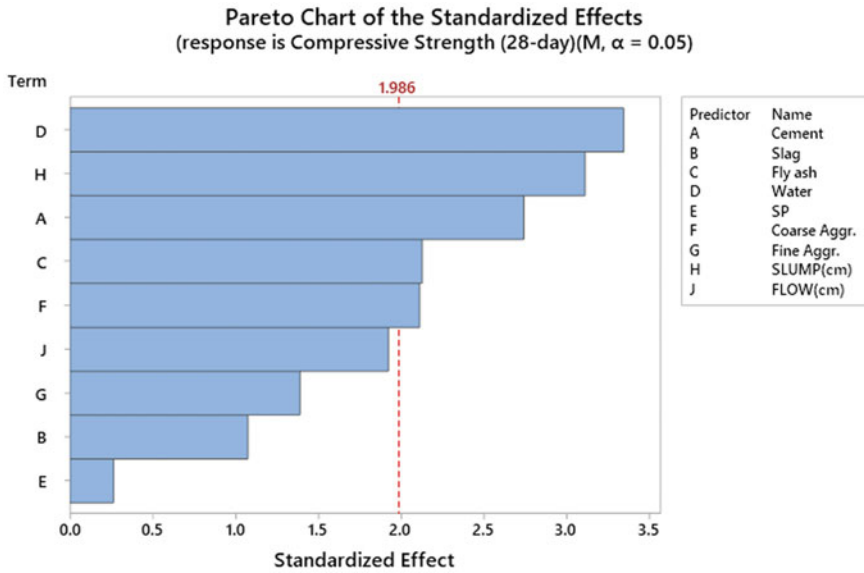


Fig. 8 Pareto chart of the standardized effects

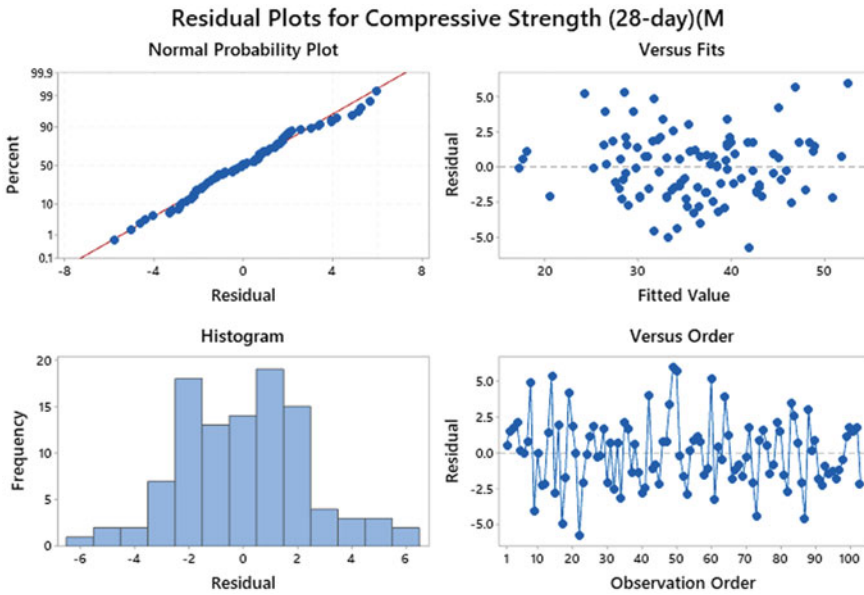


Fig. 9 Residual plots for compressive strength (28 day)

Table 5 Coefficients value for the slump

Coefficients					
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	94.8	89.2	1.06	0.290	
Cement	0.0076	0.0292	0.26	0.794	52.78
Slag	-0.0219	0.0392	-0.56	0.577	55.81
Fly ash	-0.0045	0.0293	-0.15	0.878	62.16
Water	-0.2030	0.0938	-2.17	0.033	35.66
SP	-0.293	0.166	-1.77	0.081	2.16
Coarse aggregate	-0.0296	0.0344	-0.86	0.392	92.01
Fine aggregate	-0.0244	0.0358	-0.68	0.496	51.02
Flow (cm)	0.5045	0.0261	19.36	0.000	2.08
Compressive strength (28 day) (M)	-0.399	0.128	-3.11	0.002	10.04

Table 6 Model summary and analysis of variance

Model summary					
S	R-sq	R-sq (adj)	R-sq (pred)		
3.20442	87.77%	86.59%	84.94%		
Analysis of variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	9	6855.93	761.77	74.19	0.000
Cement	1	0.70	0.70	0.07	0.794
Slag	1	3.22	3.22	0.31	0.577
Fly ash	1	0.25	0.25	0.02	0.878
Water	1	48.15	48.15	4.69	0.033
SP	1	32.00	32.00	3.12	0.081
Coarse aggregate	1	7.59	7.59	0.74	0.392
Fine aggregate	1	4.79	4.79	0.47	0.496
Flow (cm)	1	3850.01	3850.01	374.94	0.000
Compressive strength (28 day) (M)	1	99.25	99.25	9.67	0.002
Error	93	954.95	10.27		
Total	102	7810.88			

Regression equation for slump value:

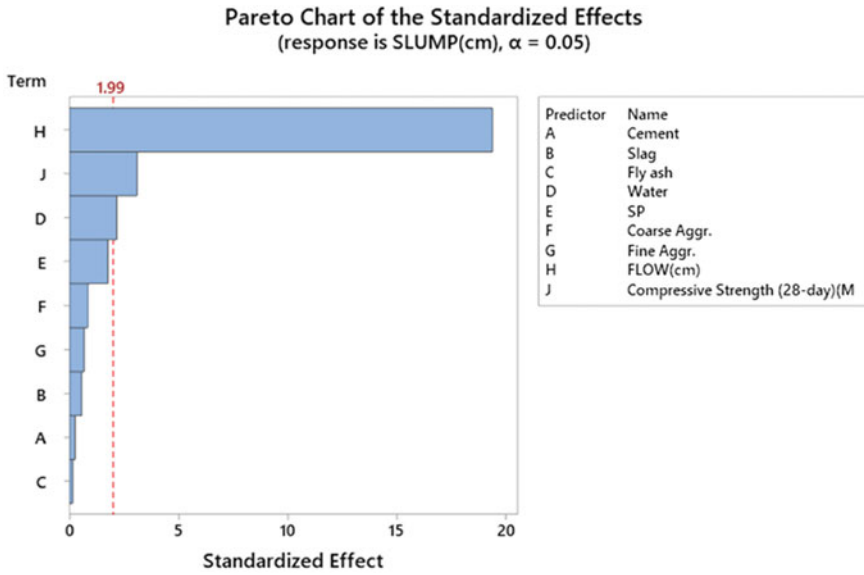


Fig. 10 Pareto chart for slump value

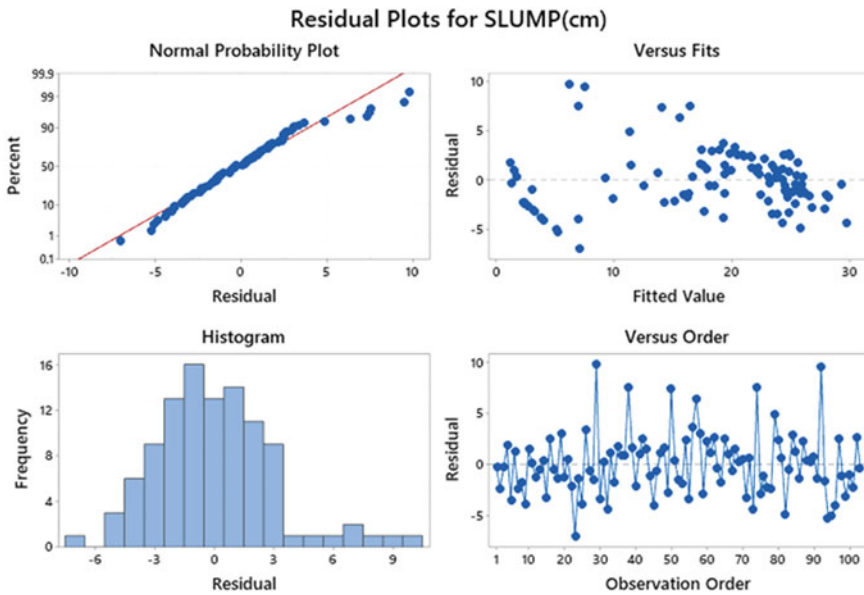


Fig. 11 Residual plots for slump value

$$\begin{aligned} \text{Slump(cm)} = & 94.8 + 0.0076 \text{ cement} - 0.0219 - 0.0045 \text{ fly ash} - 0.2030 \text{ water} \\ & - 0.23 \text{ SP} - 0.0296 \text{ Coarse Aggr.} - 0.0244 \text{ Fine Aggr.} \\ & + 0.5045 \text{ FLOW(cm)} \\ & - 0.399 \text{ Compressive Strength(28-day)} \end{aligned}$$

After training and testing the dataset, we got the ideal composition for slump value using regression method analysis and prediction for slump value. Regression analysis was performed for flow value to predict the best equation for flow value (Tables 7 and 8; Figs. 12 and 13).

Regression equation for flow value of concrete:

$$\begin{aligned} \text{FLOW(cm)} = & -175 + 0.0097 \text{ Cement} + 0.0283 \text{ Slag} + 0.0286 \text{ Fly ash} \\ & + 0.425 \text{ Water} + 0.544 \text{ SP} + 0.0515 \text{ Coarse Aggr.} \\ & + 0.0504 \text{ Fine Aggr.} + 1.5882 \text{ SLUMP(cm)} \\ & + 0.451 \text{ Compressive Strength(28-day)} \end{aligned}$$

Hence, the prediction for the ideal composition for flow value by using regression analysis was obtained.

Table 7 Coefficient for flow value

Coefficients					
Term	Coefficient	SE Coefficient	T-Value	P-Value	VIF
Constant	-175	158	-1.11	0.270	
Cement	0.0097	0.0519	0.19	0.852	52.80
Slag	0.0283	0.0696	0.41	0.685	55.90
Fly ash	0.0286	0.0519	0.55	0.583	61.97
Water	0.425	0.165	2.58	0.011	34.95
SP	0.544	0.294	1.85	0.068	2.15
Coarse aggregate	0.0515	0.0611	0.84	0.401	92.04
Fine aggregate	0.0504	0.0634	0.79	0.429	50.93
Slump (cm)	1.5882	0.0820	19.36	0.000	1.63
Compressive strength (28 day) (M)	0.451	0.235	1.92	0.058	10.66

Table 8 Model summary and analysis of variance

Model summary					
S	R-sq	R-sq (adj)	R-sq (pred)		
5.68545	90.45%	89.53%	88.42%		
Analysis of variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	9	28,476.8	3164.1	97.89	0.000
Cement	1	1.1	1.1	0.04	0.852
Slag	1	5.3	5.3	0.17	0.685
Fly ash	1	9.8	9.8	0.30	0.583
Water	1	215.7	215.7	6.67	0.011
SP	1	110.4	110.4	3.42	0.068
Coarse aggregate	1	23.0	23.0	0.71	0.401
Fine aggregate	1	20.4	20.4	0.63	0.429
Slump (cm)	1	12,119.7	12,119.7	374.94	0.000
Compressive strength (28 day) (M)	1	119.5	119.5	3.70	0.058
Error	93	3006.2	32.3		
Total	102	31,482.9			

Pareto Chart of the Standardized Effects
(response is FLOW(cm), $\alpha = 0.05$)

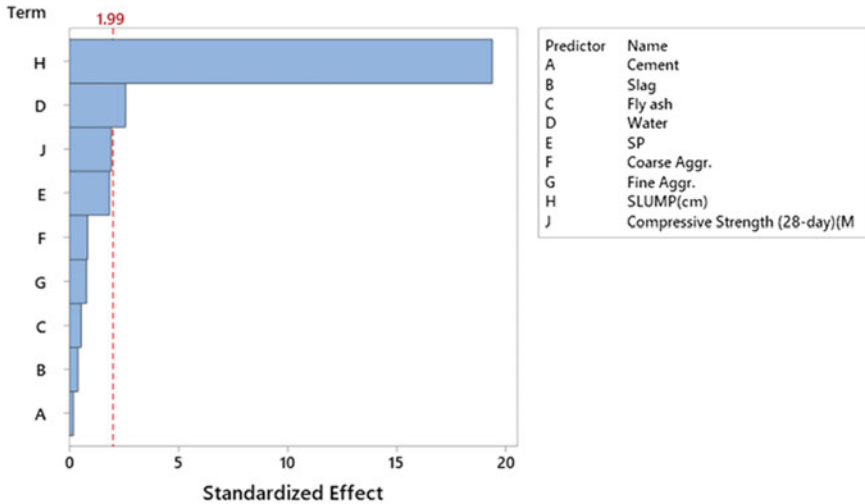


Fig. 12 Pareto chart for flow value

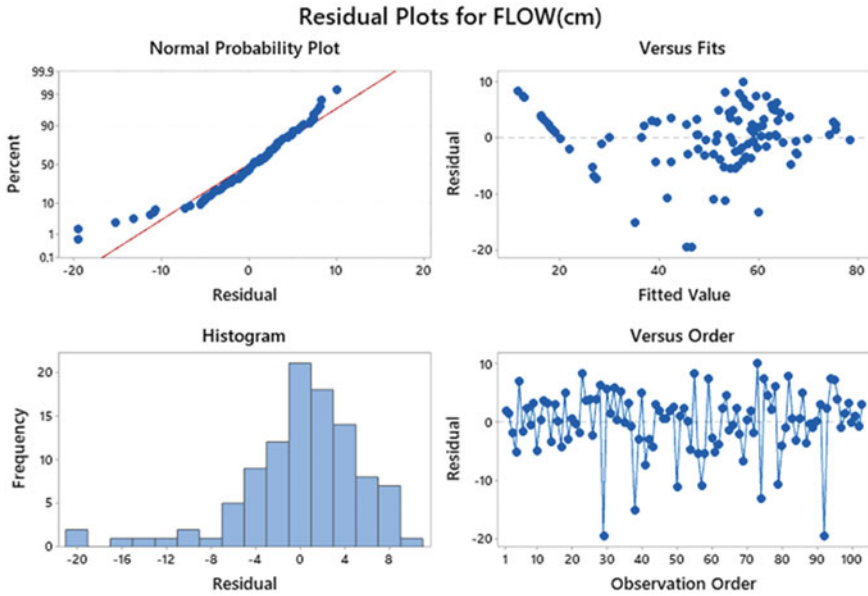


Fig. 13 Residual plots for flow value

6 Conclusion

The models were developed utilizing ANN and linear regression models from a reference dataset in this empirical analysis. The model with its respective inputs has been checked for its output with the regression equation to forecast the compressive intensity for 28 days and its workability. The plot of residual data for compressive strength is built to evaluate the fitness level after regression for various sets of input parameters, deciding whether the ordinary least square assumptions are satisfied if presumptions are met, ordinary lowest square regression results unbiased coefficient estimates with the least variance. In machine learning, the ideal value of RMSE is 0 and that of the R-squared method is 1. This model, which is about 80% accurate, can predict concrete workability and compressive strength precisely in short time. The ANN models developed will contribute to timesaving, reducing waste material and decreasing the project's overall cost. The early determination of CS of concrete and slump or flow values plays a very important role in construction.

References

1. V.K. Alilou, M. Teshnehlab, Prediction of 28-day compressive strength of concrete on the third day using artificial neural networks. *Int. J. Eng.* **3**(6), 565–576 (2010)
2. P. Singh, N.D. Shah, An experimental investigation on sustainable concrete with flyash and steel fibers. *Int. J. Civ. Eng. Technol.* **9**(6), 1131–1140 (2018)
3. P. Singh, P. Khaskil, Prediction of compressive strength of green concrete with admixtures using neural networks, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, 2020, pp. 714–717, <https://doi.org/10.1109/GUCON48875.2020.9231230>
4. B.K.R. Prasad, H. Eskandari, B.V. Venkatarama Reddy, Prediction of compressive strength of SCC and HPC with high volume fly ash using ANN. *Constr. Build. Mater.* **23**(1), 117–128 (2009)
5. P.R. Singh, A. Goel, S. Thakur, N.D. Shah, An experimental approach to investigate effect of steel fibers on tensile and flexural strength of fly ash concrete. *Int. J. Sci. Eng. Appl. Sci. (IJSEAS)* **2**(5), 384–392 (2016)
6. S. Kostic, D. Vasovic, Prediction model for compressive strength of basic concrete mixture using artificial neural networks. *Neural Comput. Appl.* **26**, 1005–1024 (2015)
7. I.B. Topcu, M. Saridemir, Prediction of compressive strength of concrete containing fly ash using artificial neural networks and fuzzy logic. *Comput. Mater. Sci.* **41**(3), 305–311 (2008)
8. P.R. Singh, N.D. Shah, Impact of coal combustion fly ash used as a binder in pavement. *Civ. Eng. Environ. Tech.* **1**, 57–60 (2014)
9. R. Cook, J. Lapeyre, H. Ma, A. Kumar, Prediction of compressive strength of concrete: critical comparison of performance of a hybrid machine learning model with standalone models. *J. Mater. Civ. Eng.* **31**(11), 04019255 (2019)
10. D. van Dao, H.-B. Ly, S.H. Trinh, T.-T. Le, B.T. Pham, Artificial intelligence approaches for prediction of compressive strength of geopolymer concrete. *Materials* **12**(6), 983 (2019)
11. H.-G. Ni, J.-Z. Wang, Prediction of compressive strength of concrete by neural networks. *Cem. Concr. Res.* **30**(8), 1245–1250 (2000)
12. A.T.A. Dantas, M.B. Leite, K. de Jesus Nagahama, Prediction of compressive strength of concrete containing construction and demolition waste using artificial neural networks. *Constr. Build. Mater.* **38**, 717–722 (2013)
13. F. Khademi, M. Akbari, S.M.M. Jamal, Prediction of compressive strength of concrete by data-driven models. *I-Manager's J. Civ. Eng.* **5**, 16 (2015)
14. Z.-H. Duan, S.-C. Kou, C.-S. Poon, Prediction of compressive strength of recycled aggregate concrete using artificial neural networks. *Constr. Build. Mater.* **40**, 1200–1206 (2013)
15. M. Saridemir, Prediction of compressive strength of concretes containing metakaolin and silica fume by artificial neural networks. *Adv. Eng. Softw.* **40**(5), 350–355 (2009)
16. A.T. Seyhan, G. Tayfur, M. Karakurt, M. Tanoğlu, Artificial neural network (ANN) prediction of compressive strength of VARTM processed polymer composites. *Comput. Mater. Sci.* **34**(1), 99–105 (2005)
17. C. Başıyigit, I. Akkurt, S. Kilincarslan, A. Beycioglu, Prediction of compressive strength of heavyweight concrete by ANN and FL models. *Neural Comput. Appl.* **19**(4), 507–513 (2010)
18. J. Noorzaei, S.J.S. Hakim, M.S. Jaafar, W.A.M. Thanoon, Development of artificial neural networks for predicting concrete compressive strength. *Int. J. Eng. Technol.* **4**(2), 141–153 (2007)
19. M. Nikoo, F.T. Moghadam, Ł. Sadowski, Prediction of concrete compressive strength by evolutionary artificial neural networks. *Adv. Mater. Sci. Eng.* (2015)
20. S.U. Khan, T. Ayub, S. Rafeeqi, Prediction of compressive strength of plain concrete confined with ferrocement using artificial neural network (ANN) and comparison with existing mathematical models. *Am. J. Civ. Eng. Architect.* **1**(1), 7–14 (2013)

Smart Attendance and Progress Management System



Lathika Krishnapillai, Sangavi Veluppillai, Adonija Akilan,
V. Naomi Saumika, K. P. Dhammika H. De Silva, and M. P. A. W. Gamage

Abstract Management of attendance may be a great burden on lecturers if done manually. This study focuses on finding an automated solution for taking attendance and keeping track of progress of a student in a smart way. The smart attendance system is generally using biometrics for identifying individuals. In this study, face recognition was considered for identification. The student's face is recognized and attendance is taken using face biometrics based on high-definition monitor camera. The images of the student are given as an input and image classification was done using CNN algorithm preventing duplicate entries for attendance. For tracking the progress of the student, the factors affecting the GPA are trained using Machine Learning algorithms. This research also aims to examine the effective progress of undergraduate students by taking past year records and find out the factors for their high and low output which will be helpful to improve their performance.

Keywords Smart attendance · Machine Learning · Convolutional Neural Networks · Face recognition

L. Krishnapillai · S. Veluppillai · A. Akilan · V. N. Saumika · K. P. D. H. De Silva ·
M. P. A. W. Gamage (✉)
Department of Information Technology, Faculty of Computing, Sri Lanka Institute of Information
Technology, Malabe, Sri Lanka
e-mail: anjalie.g@sliit.lk

L. Krishnapillai
e-mail: it17145176@my.sliit.lk

S. Veluppillai
e-mail: it17045490@my.sliit.lk

A. Akilan
e-mail: it17120180@my.sliit.lk

V. N. Saumika
e-mail: it17129954@my.sliit.lk

K. P. D. H. De Silva
e-mail: dhammika.d@sliit.lk

1 Introduction

Machine Learning (ML) is one of the essential technologies used in implementing intelligent computer applications. Based on Jeff Dean's Research, computers are now better at recognizing and analyzing images than humans. Also, it says the fault rate in humans is just 3% in computer vision [1]. So, every organization should be aware of technology and be ready to compete and sustain in the industry with better products and services provided through intelligent systems. Nowadays computers control the data and the processes better than humans and humans have to become very much depended on the accuracy of the outputs provided by such systems.

In this technological era, educational institutions are adopting to smart systems in doing their day-to-day processes and also looking at productive outputs in their activities. Traditional tasks such as marking attendance, sharing course materials, and conducting examinations have now become automated. Due to COVID-19, the latest addition to this is online delivery of lectures. Therefore, certain administrative tasks in education are now looking at better ways of improvement by using latest innovations in Information Technology.

The main aim of this paper is to investigate a possible and accurate solution for smart student attendance and also to keep track of the progress of students using the Machine Learning (ML) algorithms.

This research is an investigation of the use of face detection for attendance marking to avoid errors that occur in traditional manual attendance taking such as signing for someone else or forgetting to sign. Students' progress prediction is handled using the decision tree algorithm mainly focused on j48 to provide students with the progress information to improve their performances to get a better result.

2 Literature Review

In recent years, a number of faces recognition-based attendance management systems have been introduced in order to improve the performance of students. Jomon and Zacharia [2] have mentioned in their research paper, a system using image processing, eigenfaces, PCA, and microcontroller, based on MATLAB. This system works only with front face images but there is no suitable methods proposed for different orientations of face.

For recording the attendance through webcam, we are using the face recognition algorithms, and it helps to identify the faces of the students and compares the recognized face with database images. Once the system identifies the student, the attendance is stored in the database for further calculation, if not the system stores the captured image as a new image for future use. Also, there might be a situation where the camera may have missed any faces or captured blurry images. A 3D face recognition approach for attendance management system was proposed by Muthu Kalyani and Veera Muthu [3]. They marked attendance with monthly progress of each

student. There is need for an alternative algorithm which can enhance the recognition on oriented faces.

According to the literature, internal and external classroom factors affect students' academic performance. Internal factors can include the student's competence in language, class schedules, class size, textbooks, class test results, learning facilities, homework, the environment of the class, the complexity of the course material, teacher's role in the class, the technology used in the class and exams systems. External classroom factors include family problems, work, financial status, social and other problems that might affect the student's academic progress.

Research studies done by Naqvi [4] show that students' performance depends on many factors such as learning facilities, gender and age differences, etc., that can affect student performance. Harb and El-Shaarawi [5] found that the most important factor with a positive effect on students' performance is the student's competence in English. Many researchers have already discussed this topic in their research articles. Elton and Laurillard [6] explained that structures formed by individual lectures will affect the students. Biggs [7] argued that learning strategies differ among individuals and learning situations.

Noble et al. [8] discuss that, students' academic achievements and activities, perceptions of their coping strategies, and positive attributions, plus background characteristics are indirectly linked to composite scores of the students, through academic achievement in high school.

There have been a few relevant research papers that come close to this problem. As cited in [9], the research compares the accuracies of various Machine Learning algorithms using true positive rates, false-positive rates and thereby computing the precision and recall. The root-means-square error gives a direct way of getting accuracy. Moreover, it relayed for the classification of various modules which are being conducted in the universities.

Belachew and Gobena [10] stated in a paper that feed-forward neural networks, Bayesian classifier and Support Vector Machines were used for comparison. But throughout our research, we are going to compare the active modules along with the expected revised modules with the use of ML algorithms. Several other types of research cited in [11] and [12] use classification techniques such as decision trees and other regression techniques to predict student performance throughout the modules. Therefore, we can find the predictive results and the performance level of the students according to the modules they selected and, they can follow up the related modules with the guidance of an academic advisor.

3 Methodology

3.1 Face Detection and Attendance Update

All students must register themselves by entering the required relevant data, and then capture and store their images in the database. Faces will be identified from Webcam during each session. The identified faces will be compared to images present in the dataset. If a match is found, attendance for the respective student is marked (Fig. 1).

At the beginning of every session, an image of each student is captured using a webcam. Multiple images of an individual student will be acquired with varied gestures and angles. OpenCV is used for face detection. The imported image is transformed from RGB to grayscale because it is easy to detect a face in grayscale. After that, the trace elements were used, whereby the images were resized, clipped, blurred and sharpened if necessary. The next level is image segmentation, which is used to identify edges or segment the various items in a single image such that the classifier can identify the features and faces in the picture easily.

Face detection—For face detection Haar-like features' algorithm was used. Viola-Jones algorithm is used for finding the location of a human face in the frame before it detects the face. This is called Feature Extraction (Fig. 2).

Face recognition—The method of facial recognition can be divided into three stages—planning training data, train face recognizer, and prediction. Once the identification and preparation of the face have been done, updating the student's presence is correlated with the faces currently in the students' database. Face authentication requires a one-to-one match that matches a face image of the user against a face image of the prototype (model) whose identity is asserted. Face recognition involves a one-to-many match which compares an image of a query face again.

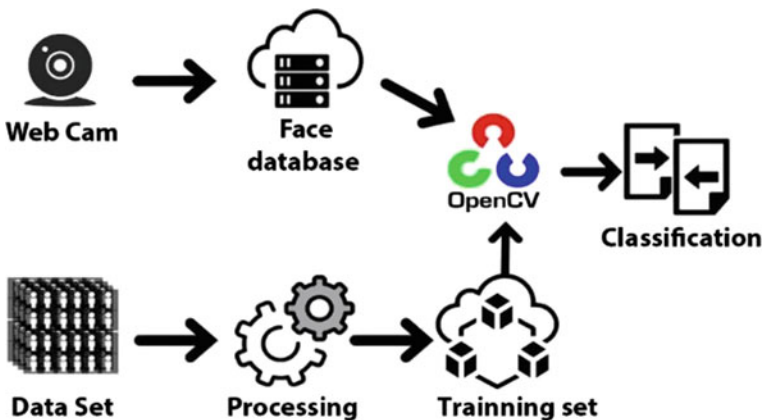


Fig. 1 System overview diagram

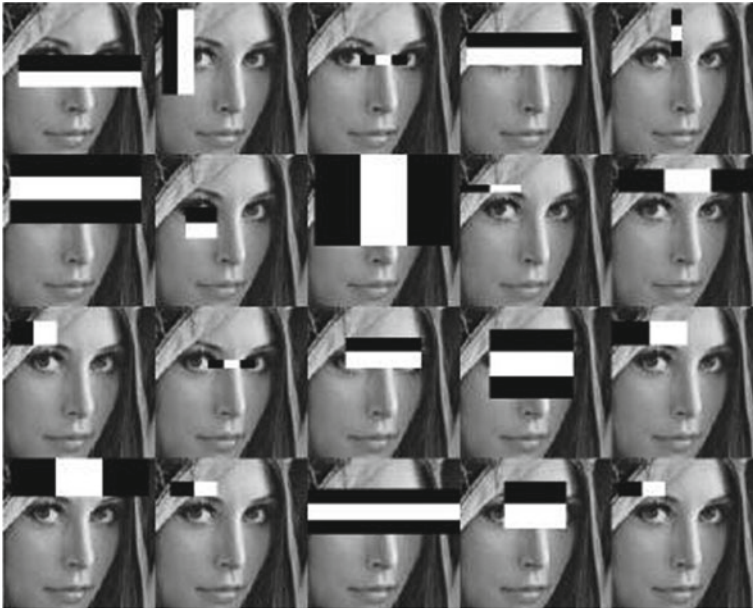


Fig. 2 Haar-like features for face detection

Attendance updating—Once the face identifies with the image that is stored in the database, it generates roll numbers of present students and returns with the system generated attendance table with the name, roll number, date, day and time with the corresponding subject id. And then passes the data into Python to automatically store the table in an Excel sheet (Figs. 3 and 4).

3.2 Face Recognition Using CNN Algorithm and Preventing the Duplicate Entry

CNN is an extremely effective methods in image classification and pattern recognition that has achieved great success. Multiple images per person are often available for training and real-time recognition is required [13]. Our strategies take advantage of the inherent correlation between face identity, smile, gender, and other face characteristics to alleviate the problem of over fitting the small training set and improve the performance of the classification [1].

The primary task of classifying images is to recognize the input image and its class description below. In this function, we need to input images which are already captured and recognized using the webcam. Here the computer sees the images quite differently. The machine sees an array of pixels instead of the images. The computer is looking for the characteristics of the base level to solve this problem. In human

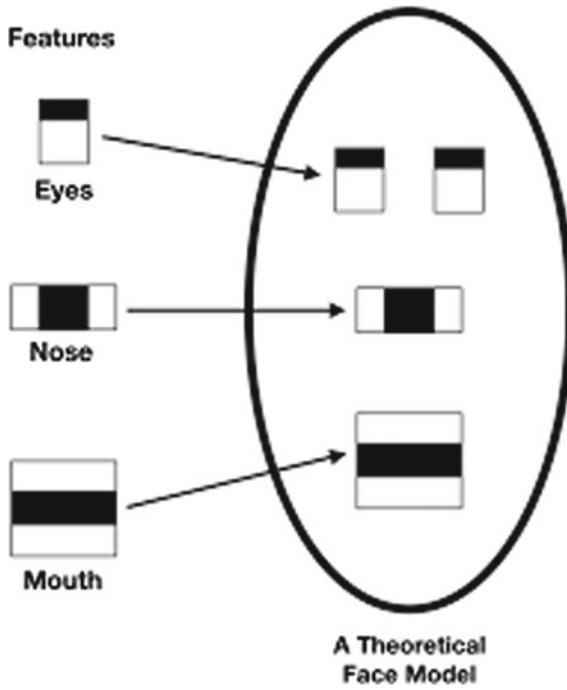


Fig. 3 Rectangular Haar-like features

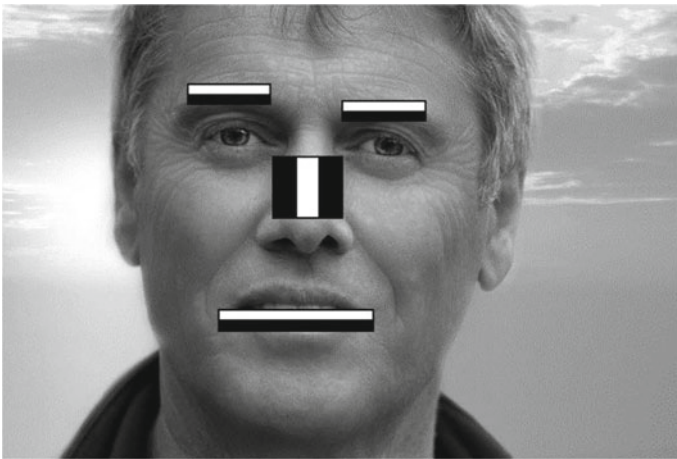


Fig. 4 Detection of the position of eyebrows, nose, and mouth



Fig. 5 Set of six images per class

comprehension, such features include, for example, eyes or ears. For the computer, boundaries or curvatures are these features. And then the computer constructs more abstract concepts through the groups of convolutional layers. A series of pooling layers and fully connected layers is passed through the image, and the output is then generated. Although we frequently believe that data points within a dataset are independent and distributed identically, this is rarely (if ever) the case when dealing with a real-world dataset. For this, we need to analyze the images and prevent duplicate entry in the database. A deep learning model requires a large amount of data, which is considered. It needs approximately 100 photographs per class for training (Fig. 5).

3.3 GPA Prediction and the Factors that Affect the GPA of a Student

Data of undergraduate students have been collected by an online survey form based on the research questions and a sample semester grade point average (GPA) and cumulative grade point average (CGPA) excel sheet was created randomly. On the

survey, student id, university, native language, study program, current semester year, advanced level scheme and results, CGP, cumulative credits, CGPA and attendance in a percentage format were collected [14, 15] (Figs. 6, 7 and 8).

To implement the classification algorithm to the collected dataset, WEKA toolkit has been used which was developed by the University of Waikato in New Zealand using JAVA [16]. It gives a collection of data mining machine learning algorithms and processing tool to do data mining and classification (Fig. 9).

To attain the principal factors which are affecting the grade point average and the relationship between the eight Semesters GPA and CGPA, we applied classification for data mining. As a classification algorithm we used C4.5(J48) algorithm in decision

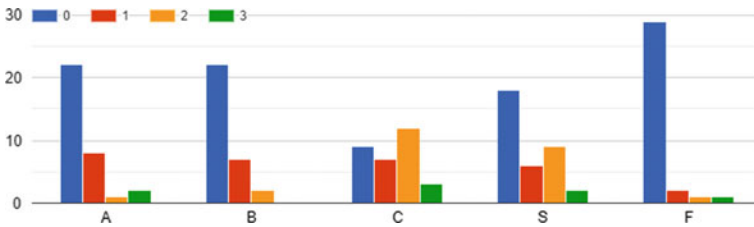


Fig. 6 Survey outcome—results of advanced level

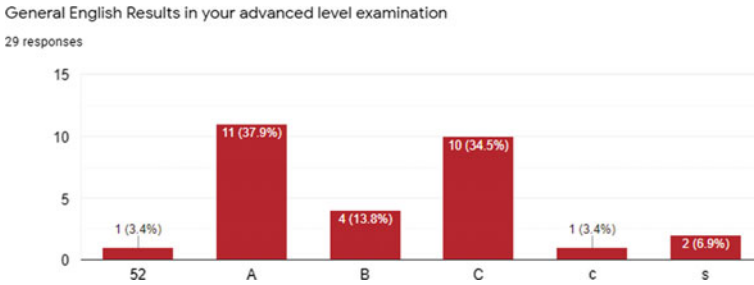


Fig. 7 Survey outcome—general English result in advanced level examination

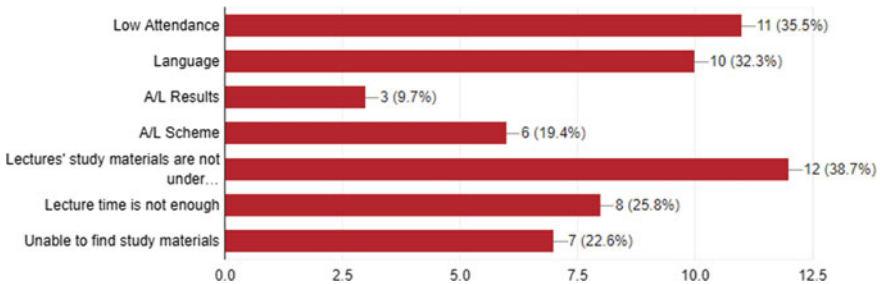


Fig. 8 Survey outcome—what are the factors affecting the GPA

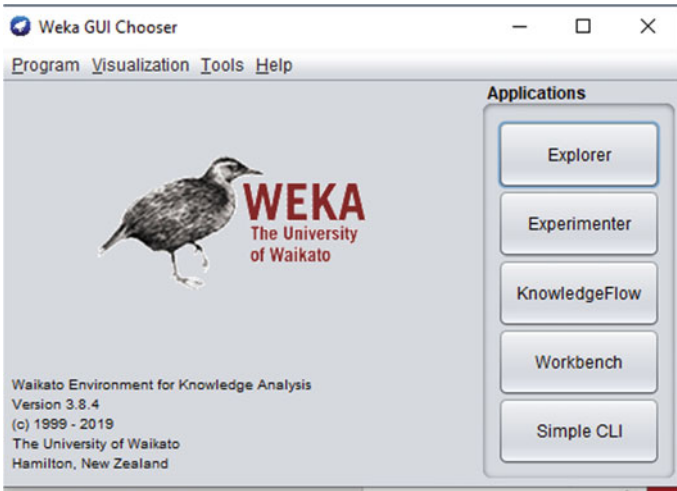


Fig. 9 WEKA toolkit UI

tree which was developed by Ross Quinlan and this represents the results in a tree form and J48 is an open-source implemented in WEKA tool kit (Fig. 10).

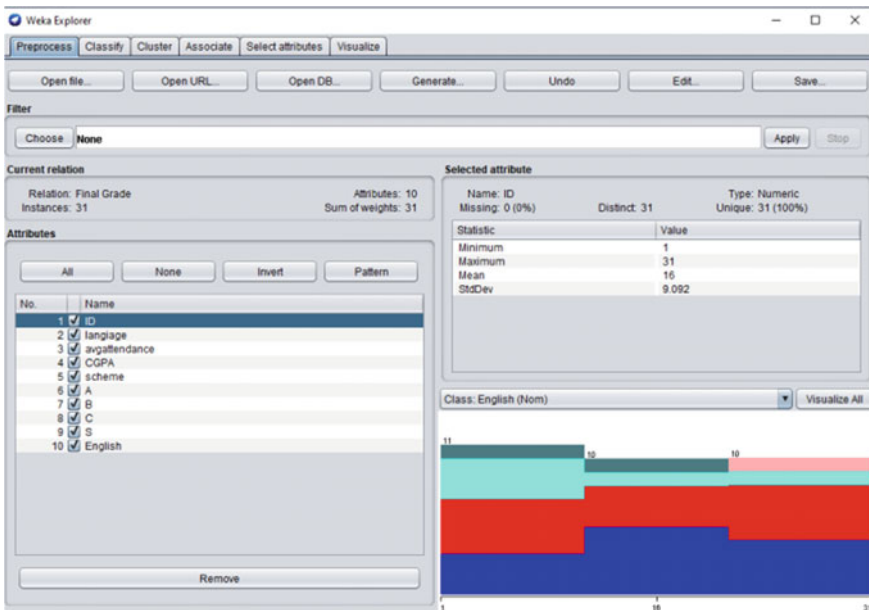


Fig. 10 WEKA classification process

Table 1 Students' level based on their GPA

0.0–2.30	Poor
2.31–2.69	Average
2.70–3.69	Good
3.70–4.00	Very good

We processed the collected data into data mining classification algorithm, after making some alteration to the dataset. First some unwanted data were removed and incorrect information such as the university, GPA importance, timestamp, and calculated an average attendance percentage based on the survey details. And to implement the J48 classification, we divided the GPA into four categories based on the GPA float value [14] (Table 1).

4 Result and Discussion

Manual performance monitoring analysis is extremely time-consuming and less accurate. This also needs expert quality, so it becomes a very costly operation. So, the purpose of this study is to concentrate on specific targets which helps to analyze and to know how students' performance levels can be changed for the specific modules by revising them using the data mining techniques after predicting and analyzing their future performance.

This research was composed of three phases namely experimentation, preparation, and testing. A number of attributes during these stages were increased step by step, for example, when only demographic information was included in the first step. In the next step, performance attributes were introduced. For each of these subsequent steps, algorithms were checked and then compared. This comparative research helped to narrow down our own application candidates. Classification of the data into binary classes, however, seems inadequate. The study's primary objective was to identify at-risk students, rather than to evaluate student performance rates. This may be more helpful to identify students in various classes according to their results (e.g., bad, average, decent, excellent, etc.). In this way, the instructors can provide each student with more adaptive feedback.

These study measurements include the gain data, gain ratio, and gain index. Algorithms used for the studies includes common decision trees (DT) (C4.5) algorithm, Naive Bayes (Bayesian networks), and fuzzy genetic algorithm (FGA). FG has two computational elements that work together: the genetic algorithm (GA) and fuzzy fitness finder (FFF).

In DT, predefined strict decisions are followed, so students at the barrier of success will be identified as a risk. So, in lecturer's point of view average-level student will also be under the care of experts. So, in final exams, a good result can be expected. In FGA, a student having a low score in some attributes has a chance to be safe due the high score attained from other attributes which make students comfortable.

4.1 Face Detection and Attendance Update

Image captured data are converted into a grayscale image to detect and recognize the face, and for the task of face detection, the initial rectangle features selected by AdaBoost are easily interpreted. The camera is fixed at a specific distance inside a classroom to capture an image of the student. The 200-feature classifier presents primary data that a boosted classifier constructed from rectangle features is an adequate technique for face detection. For the final results, the captured images need to convert into frames per second detection and recognition of the students by face images. Haar Cascade classifier is used for face detection, also LBPH algorithm was implemented for face recognition. Even though these results are compelling, still it is not sufficient for many real-world tasks. In terms of computation, this classifier is very fast. Unfortunately, the most straightforward technique for improving detection performance, adding features to the classifier, directly increases computation time.

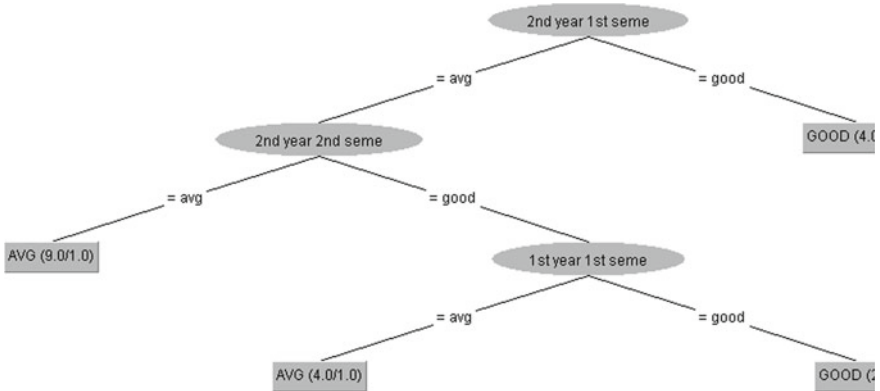
4.2 Face Recognition Using CNN Algorithm and Prevent the Duplicate Entry

In fault face detection, the expectation was the result for the recognition of face and removing the duplicate entry. Typically, duplicate images were removed from the dataset to ensure each data point (image) is represented only a single time. If there are multiple identical images in the dataset, it is difficult to calculate the attendance.

This has been tested and found that a small number of photographs work really well with this model. In order to identify potential over fitting problem, we measured how the precision depends on the number of epochs. We determined that 10 epochs are enough for successful training of the model (Table 2).

Table 2 Accuracy of the data

Student	Accuracy (%)
Student 1	75
Student 2	79
Student 3	84



4.3 GPA Prediction and the Factors that Affect the Grade Point Average of a Student

To acknowledge the research question, the relationship between semester GPA and final cumulative GPA, also the factors that affect the final grade point average of the students are archived by implementing the collected dataset into WEKA toolkit which we used for classification and data mining.

For instance, we used the graduated students’ final year CGPA and did data mining and found that the 2nd year 1st semester GPA is in the root node. If a student has a GOOD GPA in 2nd-year 1st semester, he will graduate with the GOOD CGPA (Figs. 11 and 12).

As the factors that affect the GPA of the student, factors such as native language, average attendance, advanced level result in selected schemes, English result in

```

J48 pruned tree
-----

2nd year 1st seme = avg
| 2nd year 2nd seme = avg: AVG (9.0/1.0)
| 2nd year 2nd seme = good
| | 1st year 1st seme = avg: AVG (4.0/1.0)
| | 1st year 1st seme = good: GOOD (2.0)
2nd year 1st seme = good: GOOD (4.0/1.0)

Number of Leaves : 4

Size of the tree : 7
  
```

Fig. 11 J48 classification outcome—relationship between semester GPA and cumulative GPA

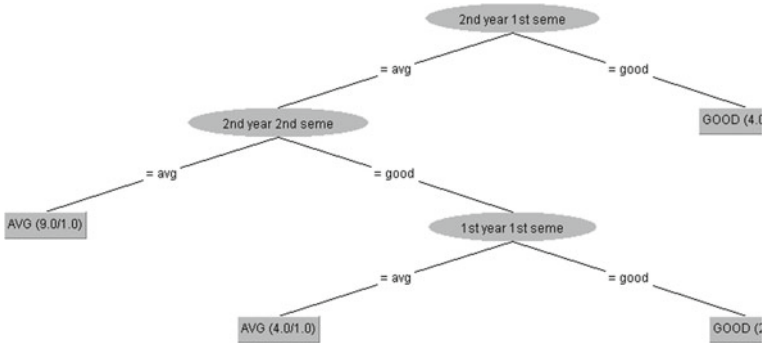


Fig. 12 J48 classification outcome of relationship between semester GPA and cumulative GPA—tree view

advanced level, and the advanced level scheme were used. Here, the English results in advanced level are taken as the root node (Figs. 13 and 14).

J48 pruned tree

```

English = C
| language = Tamil
| | B <= 0
| | | ID <= 20: AVG (5.0/1.0)
| | | ID > 20: GOOD (2.0)
| | B > 0: AVG (3.0)
| language = Sinhala: GOOD (2.0)
English = A
| C <= 0: GOOD (5.0/1.0)
| C > 0
| | S <= 0: AVG (4.0)
| | S > 0: GOOD (2.0/1.0)
English = B
| avgattendance <= 58.125: AVG (2.0)
| avgattendance > 58.125: GOOD (3.0/1.0)
English = s: AVG (2.0/1.0)
English = c: AVG (1.0)

Number of Leaves : 11

Size of the tree : 18
  
```

Fig. 13 J48 classification outcome of which factor if affecting the cumulative GPA

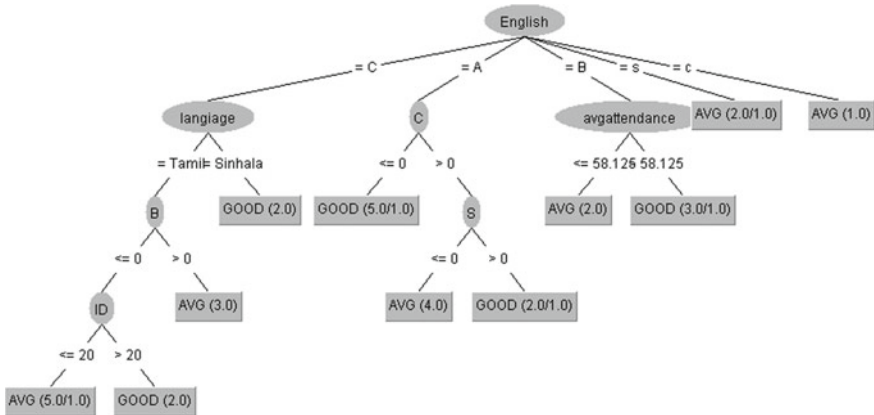
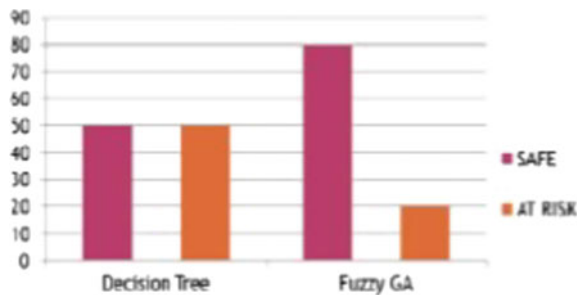


Fig. 14 J48 classification outcome of which factor if affecting the Cumulative GPA—tree view

Fig. 15 Computer networking marks



Along the graph the list of students and their status will be displayed. Following Fig. 15 shows the resulting graphs obtained after applying DT algorithm and FGA for module “Computer Networks” and big data for bachelor’s degree, respectively.

The findings of this study suggest that function engineering gives outcomes of prediction more benefit than process selection. Although feature engineering was done in a limited capacity, it made a bigger difference in the performance of predictions.

5 Conclusion

This paper presents a deep analysis and an implementation of techniques and algorithms to create a proper system that can collect student attendance and improve the academic progress of the students. CNN was used mainly to collect attendance through face recognition and making it easier to collect attendance through a web camera stationed at the entrance to a lecture hall. This will eliminate the hassle of

collection attendance manually and also avoid mistakes that happen intentionally or unintentionally. It will also save lot of time in maintaining and keeping track of the attendance. The data collected will help to analyze the student's performances through a predication system which combines more attributes which will determine the students' GPA.

References

1. The Future of Machine Learning. Medium (2020). [Online]. Available: <https://towardsdatascience.com/the-future-of-machine-learning-ce0a9dc18cb8>. Accessed: 17 July 2020
2. J. Joseph, K.P. Zacharia, Automated attendance management system using face recognition. *Int. J. Sci. Res. (IJSR)* **2**(11), 327–330 (2013)
3. K. Muthu Kalyani, A. Veera Muthu, Smart application for AMS using face recognition. *Comput. Sci. Eng.* **3**(2), 13 (2013)
4. S. Naqvi, Factors affecting students' performance a case of private colleges. *Bangladesh e-J. Sociol.* **3** (2006)
5. N. Harb, A. El-Shaarawi, Factors Affecting Students' Performance (2006)
6. L. Elton, D. Laurillard, Trends in research on student learning. *Stud. High. Educ.* **4**(1), 87–102 (1979)
7. J. Biggs, Individual and group differences in study processes. *Br. J. Educ. Psychol.* **48**(3), 266–279 (1978)
8. K. Noble, M. Wolmetz, L. Ochs, M. Farah, B. Mc Candliss, Brain-behavior relationships in reading acquisition are modulated by socioeconomic factors. *Dev. Sci.* **9**(6) (2006)
9. S. Mandal, S. Biswas, V.E. Balas, R.N. Shaw, A. Ghosh, Motion prediction for autonomous vehicles from Lyft dataset using deep learning, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, 30–31 Oct 2020, pp. 768–773. <https://doi.org/10.1109/ICCCA49541.2020.9250790>
10. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering*, ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh. *Lecture Notes in Electrical Engineering*, vol. 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
11. https://www.researchgate.net/publication/326261079_Face_detection_system_for_attendance_class'_student
12. E. Belachew, F. Gobena, Student performance prediction model using machine learning approach: the case of Wolkite University. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **7**(2), 46–50 (2017)
13. P. Viola, M.J. Jones, Robust real-time face detection. *Int. J. Comput. Vision* **57**(2), 137–154 (2004)
14. S. Lawrence, C. Giles, A.C. Tsoi, A. Back, Face recognition: a convolutional neural-network approach. *IEEE Trans. Neural Networks* **8**(1), 98–113 (1997)
15. K. Zhang, L. Tan, Z. Li, Y. Qiao, Gender and smile classification using deep convolutional neural networks, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops* (2016), pp. 34–38
16. M. Al-Barrak, M. Al-Razgan, Predicting students final GPA using decision trees: a case study. *Int. J. Inf. Educ. Technol.* **6**(7), 528–533 (2016)

Performance and Parametric Analysis of IoT's Motes with Different Network Topologies



Rajkumar Gaur  and Shiva Prakash 

Abstract Nowadays, IoT applications provide the effortless application by connecting more physical devices to the Internet. However, it is more new challenges as power consumption, low energy, the transmission of data rate, motes' capacity. So, IoT is more suitable in different areas like e-health, e-cash, e-hospital, and e-commerce et al. This paper analyzes and calculates motes power consumption in different topological based, which show the base of the different comparison on motes parameters. The sky motes deploy in different scenarios as statics and dynamics. These scenarios create the networks and evaluate improvement network performance in a different application of other papers—the performance count as sky motes power, duty cycle, energy efficiency, and security of applications. IoT application evaluations' performance clearly shows our analysis's effectiveness in improving fast response and decreased energy consumptions of applications. Then, discuss and evaluate the strategies used to decrease energy consumptions. In Contiki and Cooja, the motes are the wireless sensor network. The motes are using in the Contiki and Cooja simulator as a sensor, which is deployed in the testing of different applications of IoT because the Internet of things (IoT) denotes the static and dynamic connectivity of physical devices with limited resources that support of Internet in different types of infrastructure.

Keywords Internet of things · Wireless sensor networks · Motes · Energy efficiency · Duty cycles

1 Introduction

In this time, the network wirelessly connected to devices with the help of the sensor. The structure is the sensor, machine, terminals, mobile devices, laptops, etc. That communicates with each other, so they generate data according to your requirements, used by the wireless sensor network. The powerful main idea is the high impact on a different aspect of user's per day life and behavior, in this context, e-health, e-

R. Gaur (✉) · S. Prakash
Madan Mohan Malaviya University of Technology, Gorakhpur, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_61

787

ticketing, e-business, etc. [11]. In everyday things, for example, in food packages, furniture, paper documents, and more; however, the IoT idea poses different new problems concerning the networking aspects. Low resources will characterize the things composing the IoT in both computation and energy capacity [19]. This work focuses on implementing and analyzing motes with different topology types and various types of motes in a simulator [14, 16]. The research and the challenges are coming from implementing the network and computation of motes [34]. They propose an implementation, extending the standard Contiki's network stack, rime that allows an efficient realization of motes operation while maintaining a limited impact on data communications[30]. Finally, they validate our implementation, utilizing a set of experiments running on a real Contiki and Cooja simulator [20].

The motes design based on the following low duty cycle principle: the node is asleep for most of the time, wakes up quickly on an event, processes, and returns to sleep. For the lowest power consumption [18], the standby current and wake-up time take minimal, since the active portion of a sensor network application is typically tiny. Motes offer more than just low power operation through its integrated design. The integration of programming, communication, storage, and sensing allows researchers to utilize more functionality and develop more robust systems. So, the types of motes behaving as a sensor, each sensor performs some tasks implemented in this paper. In IoT, the protocol is chosen based on low power and lossy network as RPL and 6LoWPAN [27, 33]. With the help of motes used in Contiki and Cooja, we will try to make the best possible network planning or topology and maximum throughput of sensors for establishing communication with each other [14].

The rest of the paper is organized as follows: In section second, highlight the main feature of IoT protocols used in IoT applications. In the third section, detail the motes (nodes or sensors) of applications and properties of motes. Section fourth describes the simulation results and analysis. Finally, section fifth is of conclusions with references.

2 IoT Protocols

Most IoT devices connect to the Internet via IP, non-IP, RFID, Bluetooth. When devices to device communicate with each other, they are connected to locally or globally. If the motes connect globally, use the Internet IP address mode via a router or other devices because these devices take more power, memory but the non-IP addressing takes less power and memory because every device is a limitation. As far as the IoT communication protocols concerned, a mix of both IP and non-IP networks. The motes depend on different applications like hospital, industries, railway, Indian force services, road light automation, agriculture, and mostly used in disaster management today to save life animals and humans [2].

IoT communication protocols are modes of communication that protect and ensure minimum security transferring to the data between connected devices. In our paper, use Routing Protocol for Low Power and Lossy Networks (RPL) and IPv6 over

Low-power Wireless Personal Area Networks (6LoWPAN) protocols which use in devices communications such as the RPL which makes connectivity of upper layers and 6LoWPAN adaption layer used the IEEE 802.15.4 MAC, IEEE 802.15.4 physical layers for connection establishment and IPv6 over Low Power and Wireless Private Area Network (6LoWPAN)[3, 10]. So, now describe, first, the RPL protocol after that some 6LoWPAN protocol functionalities essential for IoT application [26].

2.1 Routing Protocol for Low Power and Lossy Network(RPL)

RPL is the IPv6 type of protocol that comes under the low power and lossy network, this protocol some limitations as terms of energy, processing, and memory communicate with different nodes [2, 16, 17]. This protocol creates a multi-hop hierarchical topology for nodes. Each node can send data to its central node, which puts it forwards upwards until it reaches the communicates or gateway node. Similarly, the sender node can send a unicast message to target a specific node in its network. RPL successfully and efficiently manages data routing for nodes with restricted resources; it provides an operational framework that ensures directional connectivity, reliability, flexibility, robustness, and scalability [21]. The key features of RPL come from its efficient hierarchy, the use of timers to minimize control messages, and the flexibility of the objective function [19]. This protocol communicates for three types to cluster nodes which are motes and organized as a destination-oriented directed acyclic graph (DODAG) which shown in Fig. 1 RPL corresponds to the motes control messages are four types and follow as [1, 18]:

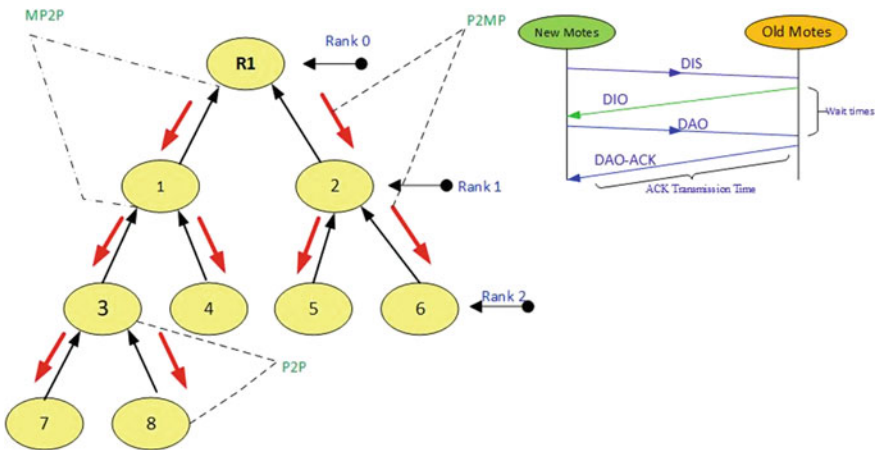


Fig. 1 Routing protocol information flow

Table 1 RPL control messages and their code message type and this will change

S. No.	Code	Description
1	0x00	DODAG information solicitation
2	0x01	DODAG information object
3	0x02	Destination advertisement object
4	0x03	Destination advertisement object acknowledgement

1. DODAG information object (DIO)
2. Destination advertisement object (DAO)
3. DODAG information solicitation (DIS)
4. DAO-ACK.

All these are communication types of motes show in Fig. 1, and the traffic pattern could be point to point (P2P), point 2 many points (P2MP), many points 2 points (MP2P). In RPL protocol, some useful codes follow in transmitting times to motes, as shown in Table 1 as an example. These codes do not fix; they always change according to communication [17, 25].

So RPL protocol, Low Power and Lossy Networks (LLN) that take fewer resources. The RPL protocols different types use a technique like processing power, battery life, memoryless, interconnect is characterized by unstable links with high loss rates, low data rate, and low packet delivery rates [4, 24]. Table 1 shows the running code which communicates motes to motes to display the information result is shown in Fig. 2 in the mote output window [10, 15].

2.2 IPv6 Over Low Power and Wireless Private Area Network (6LoWPAN)

6LoWPAN is a standard protocol imperious for the efficient usage of IPV6 over low power low rate wireless sensor networks via an adaption layer [1, 5, 13]. Therefore, the name 6LoWPAN has been derived by the combination of IPV6 and low power sensor networks [34]. The IEEE 802.15.4 standard offers the solution by introducing data rates of 20 to 250 kb/s depending on which frequency channel is using. During communications, specific protocols lay much overhead IPV6, which has much longer headers that can lodge the rest of the available bandwidth, which turns out to be a real setback while imposing Internet protocol (IP) to wireless sensor network (WSN)[20]. To resolve this problem, 6LoWPAN launched in which, with the help

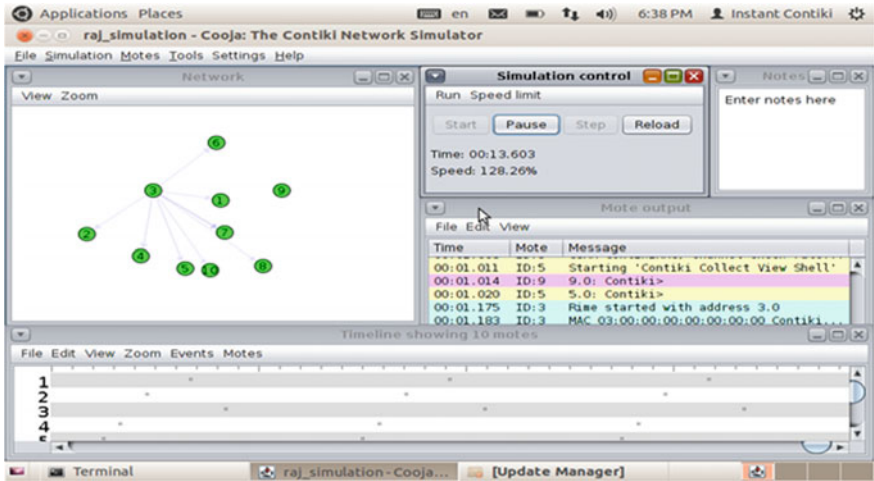


Fig. 2 Sky motes communication and desired output in different windows

Table 2 6LoWPAN dispatch code which change, its does not fix

S. No	Bit pattern	Short code	Description
1	00 xxxxxx	NALP	Not a LoWPAN Packet
2	01 000001	IPv6	Uncompressed IPv6 Addresss
3	01 000010	LoWPAN HC1	HC1 Compressed IPv6 Header
4	01 010000	LoWPAN BC0	BC0 Broadcast Header
5	01 111111	ECS	Additional Dispatch Octets follow
6	10 xxxxxx	Mesh	Mesh Routing Header
7	11 000xxx	FRAG1	Fragmentation Header (First)
8	11 100xxx	FRAGN	Fragmentation Header (Subsequent)

of an adaption layer [8, 9], header compression and fragmentation made possible, making the IPV6 overhead reduced. The 6LoWPAN protocol uses some dispatch code when transferring the data into the network to communicate the mote to motes [8, 9, 11]. Table 2 shows the codes. This code does not fix for each communication.

Table 2 shows the bit pattern with shortcodes when communicating layers to layers. Each bit code some short information; this information is showing in detail by description. Sniffer tools like Wireshark also analyze these. To review the goals of our monitoring system, our aim to maintaining a highly reliable IoT network structure via Contiki and Cooja by with different scenarios [12, 20, 32]:

1. All the motes are the statics.
2. A sink mote is static, and other motes are dynamics.
3. All the sensors and sink are dynamics.

So, all the above using the techniques found the different results according to the deployment of nodes in networks.

- Proactively and proficiently validating the correct operation of motes and relations concerning energy, temperature, ETX, and average radio duty cycle.
- Collecting, aggregating, and filtering real-time data from motes to motes beacon interval.
- We are detecting and localizing abnormal events or faults.

We are adapting to dynamic and real-time changes in the network state by network graph and network map.

3 Features of Mote's

Our analysis of the motes focuses on the Contiki platform, this motes power consumption, and motes' features describe further research on sensor networks [3, 7, 29]. The power consumption like radio traffic, motes position, and radio environment module is just analyzing different features. The power consumption of the sky mote for various operations compared to the existing additional motes [6, 16, 31]. The sky motes are automatically put to sleep when there is no active processing entering the lowest power mode when the system idles. Each mote has different features like mote IDs, addresses, radio traffic, positions, and radio environment, as shown in Table 3.

Our experiment focuses on the sky motes, which have different characteristics. These use the cc2420RF chip, MSP430F1611 model no, and microcontroller. These motes work on the IEEE802.15.4 WSN. Sky motes memory size is 48kb ROM and integrated with chips that convert the analog signal to digital and digital to analog and measure humidity and temperature light via sensors. Its interoperability with many different ieee802.15.4 devices and sky motes also used low power and fast wake-up from sleep devices. Sky motes follow the mesh, tree, and star topology. The node type may be shared between other nodes and determines features common to all these sky motes. For example, the same type's mote runs the same program code on the same simulated hardware peripherals. The same kind of motes are initializing with the same data memory, except for the mote ID. However, during execution, the motes' data memories will eventually differ after reacting to external stimuli. Table 3 shows the description of the motes output information with different features.

Table 3 Sky mote’s features

Fields	Features
Topological graphs	Sensor Map, Network Graph,
Sensor-Related Plots	Humidity Sensor—Relative Humidity. Temperature Sensor—Average Temperature and Temperature. Battery Sensor—Battery Indicator and Battery Voltage Light Sensor—Light 1 and Light 2
Network Metrics Related Plots	Neighbors, Beacon Interval, Network Hops, Router Metric (Over Time) Over Time, Instantaneous, Per Node, Average
ETX (Over Time)	Expected number of transmissions
Next Hop (Over Time)	Mechanism of adding next nodes after time out
Latency	Calculate the latency (the round-trip time from motes to motes) of a packet in a wireless link. [Latency = Data received time at sink node—Data send time from any node]
Lost Packets (Over Time)	Every network will encounter like packet loss, from time to time
Received Packets	Over Time, Per Node, Every 5 min
Over Related Plots	Average Power, Instantaneous Power, Power History, Radio Duty Cycle

3.1 Topology of Motes

The analysis of motes based on different features which system is the work on the other way as the Contiki and Cooja emulator with the various models used, which are design for the IoT environment. The sensor node is a small device used to capture sensory information, processing, and communicating this information to the node. It supports both multicast and unicast messages. According to the motes, there are currently many numbers of devices connected to the IoT environment. The IoT protocol stack supports the communication between the devices connected to the system. According to the emulator that provides many types of radio medium with start-up delay as 1000 for each medium and random no of seeds are also the same as 123,456. So, every mote’s have a different radio medium as unit disk graph medium(UDGN): distance loss, UDGn content loss, directed graph radio medium, no radio traffics and multi-path ray-tracer medium. So, Table 4 shows the different topology construction via Contiki and Cooja.

Table 4 Create motes type: Compiler Contiki for sky (topology construction)

S. no	No of motes	Positioning	Position interval
1	10	Random positioning	X 0 100 Y 0 100 Z 0 0
2	10	Linear positioning	X 0 100 Y 0 100 Z 0 0
3	10	Ellipse positioning	X 0 100 Y 0 100 Z 0 0
4	10	Manual positioning	X 0 100 Y 0 100 Z 0 0

4 Simulation Results and Analysis

During simulating various networks' performance, the Cooja simulator runs over Contiki OS. Evaluate protocols work Cooja, which is a Java-based simulator. Cooja also operates as an emulator, where code directly be installed on sensor nodes (for example, Sky mote), supporting both Windows and Linux platforms. The topology node arrangements are considering testing the detection capability of IoTs [17]. The simulation performs to analyzes the performance of the IoT various motes with different network topology and what the effect arises during communicating with each other, the communication to motes to motes via protocols, the protocol RPL with 6LowPAN [15]. In the simulation times, different scenarios construct to perform the protocol assessment. As a result, the procedures involved a different number of motes for each simulation. The motes numbers vary as 10, 20, with is other mediums of radio and motes positioning. These motes were static and no mobility implementation. The following section builds up the result and analysis performed from the simulation. The different parameters like sensor temperature, network graph, the average power consumption of sensors concerning CPU, radio transmit, radio listing and low power mode (LPM), real-time (RT) metric, expected transmission count (ETX), and power values are analyzed [18, 19].

4.1 Network Topology Construction

Figure 2 shows the Contiki 2.7 and Cooja emulator to implement the IoT application for various networks scenarios. In the upper left corner window show network, another no. of motes deployed with many types topology. The right side window

Table 5 Simulation parameters of motes

Num. of motes	10
Node type	Sky
Operating system	Contiki 2.7
Simulator	Contiki-Cooja
Routing protocol	RPL and 6LoWPAN
Traffic generated	UDP
Radio medium	Unit Disk Graph Medium: Distance Loss
Transmission range	25 m
Interference range	50 m
Simulation duration	1–20 min
Number of simulations	4 times
Confidence interval	95
Legitimate VN updates	5 updates per simulation
Per node traffic rate	1 packet per minute
Topology dimensions	20 × 20 m

shows the simulation control, the next right, the notes of algorithms, or introductory notes, and just the middle window shows the motes output. The last lower part of the window shows a timeline of motes with various colors concerning motes. In the timeline of ten motes in blue color indicates transmitting, green means receiving, and red indicates collision. The color length shows the data packet transmission time, the space between two color shows turnaround times, the blue color length shows ACK transmission times, blue and green color space known as wait times. After the simulation, more results were found via collect view only random positioning topology, and other topology results show in Tables 6, 7, 8, and 9. These results are based on various parameters, which shown in Table 5.

4.2 Network Average Power Consumption

The power of some period usually intensity varies in some way that repeats itself over time then need the average power one single period. In the above figure, three present simulation results were obtained as the CPU, radio listen, radio transmit, and low power mode (LPM) average power consumption of motes. The simulation parameters were used for topology construction as random positioning and Table 5 as simulation parameters. In the simulation, each moderate node power consumes to communicate with each other's [20, 21]. The colors combined as red show the CPU power, blue radio listens, green radio transmit, yellow show LPM, and all these data values are shown in Table 6, which as for random positioning topology (Fig. 3).

Table 6 Power consumption of Sky motes to deploy in random positioning

Nodes	Hops	Rtmetric	ETX	CPU power	LPM	Listen power	T.power	Power	LDC	TDC
1.0	1.000	8.000	1.000	0.385	0.152	0.486	0.108	1.131	0.801	0.201
2.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3.0	1.000	8.273	1.034	0.390	0.482	0.482	0.082	1.105	0.812	0.142
4.0	1.000	18.909	2.364	0.365	0.409	0.409	0.092	1.018	0.621	0.174
5.0	2.000	61.000	0.909	0.377	0.454	0.454	0.100	1.082	0.703	0.187
6.0	2.455	30.545	2.364	0.369	0.418	0.418	0.095	1.034	0.623	0.178
7.0	3.000	45.818	2.364	0.372	0.431	0.431	0.094	1.049	0.702	0.171
8.0	2.000	28.636	2.580	0.375	0.413	0.413	0.080	1.020	0.654	0.150
9.0	2.000	61.000	0.909	0.362	0.411	0.411	0.090	1.016	0.613	0.170
10.0	1.000	18.909	2.364	0.378	0.439	0.439	0.067	1.036	0.710	0.119
Average	1.546	28.109	1.589	0.337	0.361	0.394	0.081	0.949	0.624	0.149

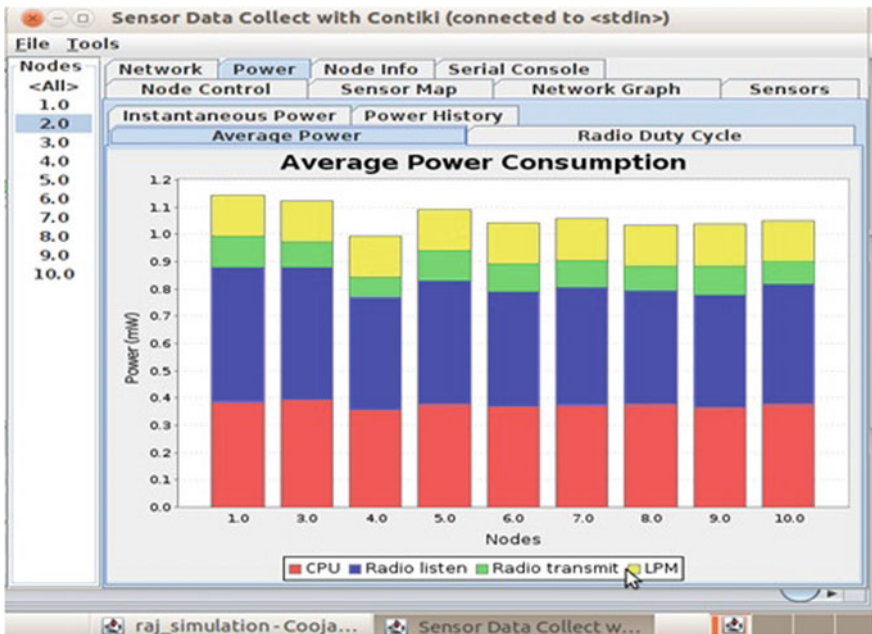


Fig. 3 Sky motes average power consumption

4.3 Network Instantaneous Power Consumption

Instantaneous power would be the amount of power delivered or required for a short time, not a big-time, as motes take instance power when they are going to sleep to active mode. Figure 4 presents an instance power consumption of motes the CPU power, radio listen, radio transmit, and LPM. In the topology construction, as random

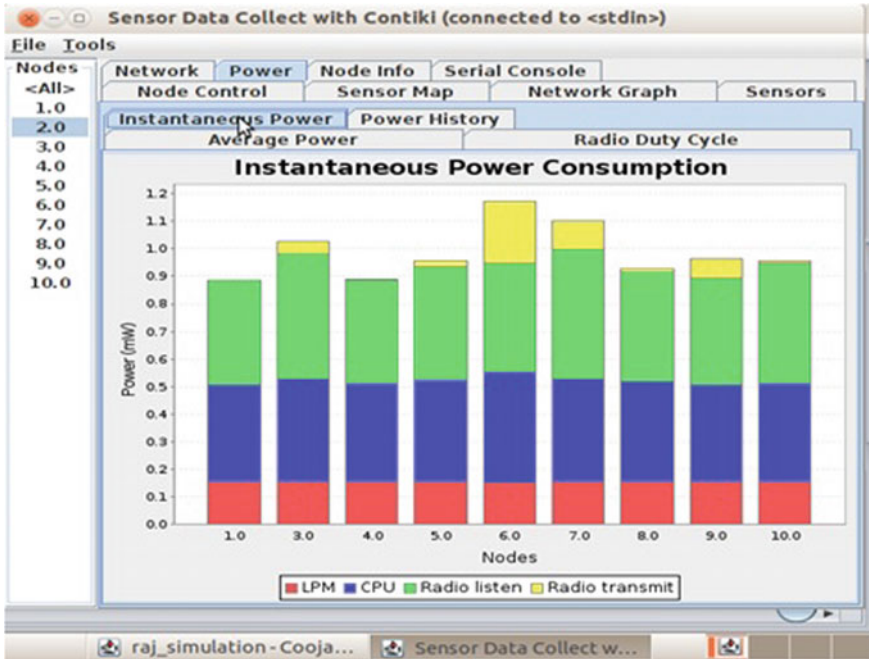


Fig. 4 Sky motes instantaneous power consumption

positioning and parameters take Table 5. In the simulation, each node communicates with each other how much instantaneous power consumption, which shows various colors combination red low power, blue CPU power, green radio listen, yellow radio transmit, and respectively data values in Table 6.

4.4 Network Graph and Sensor Map

The network graph of different topology construction shows the visualization of devices on a network, their interrelationship, and the simulator's transport layer.

Figure 5 presents a network map of devices that random topology construction in the figure. If the topology changed the node's communication and data transfer, all information would adjust according to the network. In this, simulation is each node how to communicate with each other. In Fig. 6, the current sensor map of the IoT network. The map shows the motes 2 has more nodes connect for communication in the system, so they are following DODAG and full duty cycle, max CPU processing, etc., to calculated by the average for each node by table 6 [22, 25]. Table 6 shows all the information about each mote. Each mote how many average radios duty cycles

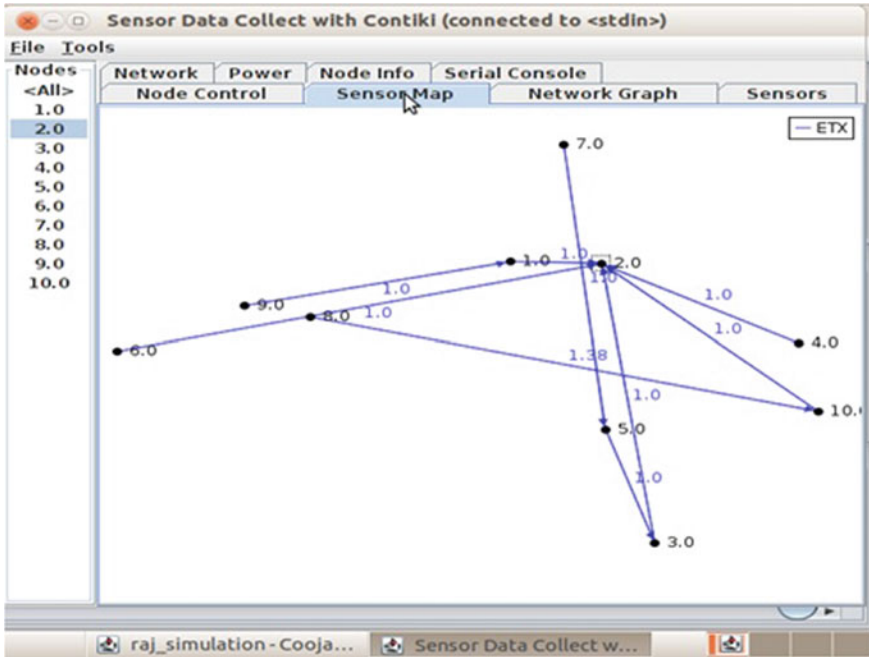


Fig. 5 Sky notes network graph

for motes with radio listeners and radio transmit and the sensor’s temperature is constant. The simulation results show in Table 6 [30].

The simulation parameters are used in Table 5 and random topology construction is shown in Fig. 6. In the simulation, each node communicates with each other concerning various fields like nodes, received times, lost packet, hops, routing metric (Rtmetric), EXT, Beacon interval, reboots, and others view in Table 6 (details description of all nodes). The measured performance metrics are packet delivery ratio (PDR), power consumption, percentage of duty cycle, etc. Tables 7, 8, 9 are using in various topologies like linear, ellipse, and manual positioning. These tables show the information of motes all types of topology construction of IoT networks.

4.5 Average Radio Duty Cycle

The average duty cycle is the proportion of times during which motes, devices, or network components operate. In Fig. 7, the bar chart shows radio listening values, which is indicated by the red colors and blue colors of all motes show the radio transmits. Each mote’s values are shown in Table 6 as random positioning, and other tables like 7,8,9 offer different topology construction data of motes.

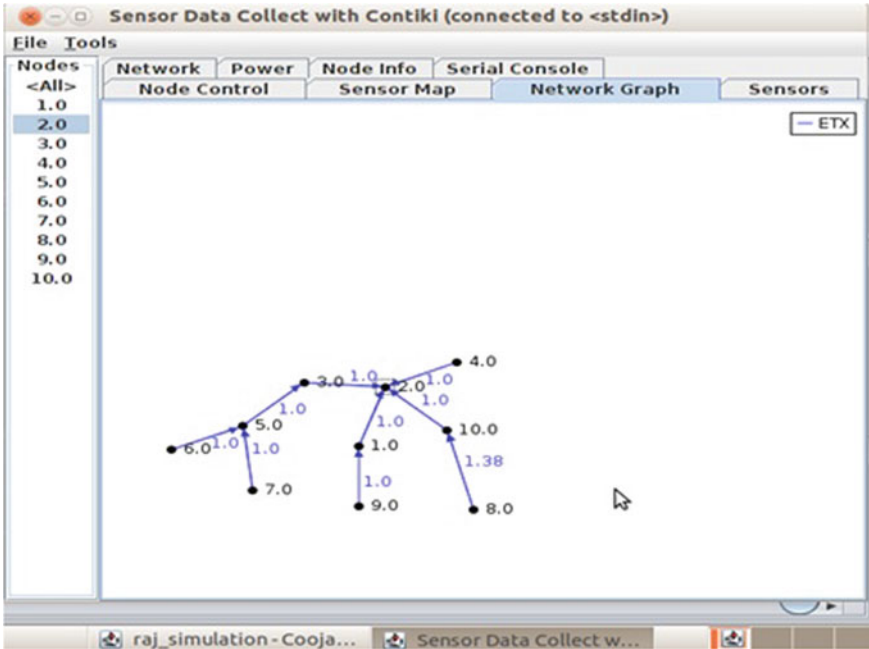


Fig. 6 Sky motes sensors map

Table 7 Details transmission and power consumption of Sky motes to deploy in linear positioning

Nodes	Hops	Rtmetric	ETX	CPU power	LPM	Listen power	T.power	Power	LDC	TDC
1.0	1.000	8.000	1.000	0.365	0.142	0.476	0.106	1.130	0.803	0.200
2.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3.0	1.000	8.173	1.030	0.380	0.470	0.478	0.081	1.104	0.815	0.145
4.0	1.000	18.509	2.264	0.335	0.410	0.405	0.090	1.016	0.620	0.172
5.0	2.462	60.000	0.709	0.357	0.435	0.445	0.101	1.084	0.708	0.185
6.0	2.000	31.245	2.264	0.358	0.411	0.415	0.098	1.037	0.620	0.175
7.0	3.000	46.818	2.354	0.369	0.429	0.429	0.096	1.050	0.705	0.170
8.0	2.000	23.636	2.570	0.368	0.418	0.420	0.082	1.022	0.652	0.152
9.0	2.000	59.000	0.908	0.356	0.416	0.415	0.089	1.019	0.611	0.172
10.0	1.000	17.900	2.344	0.388	0.437	0.440	0.065	1.035	0.714	0.117
Average	1.717	27.328	1.544	0.328	0.357	0.392	0.081	0.950	0.625	0.149

Table 8 Details transmission and power consumption of Sky motes to deploy in ellipse positioning

Nodes	Hops	Rtmetric	ETX	CPU power	LPM	Listen power	T.power	Power	LDC	TDC
1.0	1.000	8.001	1.000	0.384	0.150	0.484	0.106	1.130	0.803	0.200
2.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3.0	1.000	8.272	1.033	0.391	0.480	0.483	0.083	1.103	0.815	0.141
4.0	1.000	18.908	2.362	0.364	0.404	0.410	0.094	1.016	0.620	0.173
5.0	2.000	61.001	0.907	0.376	0.452	0.452	0.101	1.083	0.704	0.188
6.0	2.455	30.544	2.363	0.368	0.420	0.419	0.096	1.034	0.625	0.179
7.0	3.000	45.819	2.363	0.370	0.435	0.433	0.095	1.047	0.705	0.170
8.0	2.000	28.637	2.582	0.372	0.415	0.415	0.081	1.021	0.655	0.151
9.0	2.000	61.001	0.908	0.360	0.414	0.413	0.092	1.015	0.614	0.173
10.0	1.000	18.905	2.363	0.376	0.440	0.440	0.065	1.034	0.711	0.120
Average	1.546	28.109	1.588	0.336	0.361	0.395	0.081	0.948	0.625	0.150

Table 9 Details transmission and power consumption of Sky motes to deploy in manual positioning

Nodes	Hops	Rtmetric	ETX	CPU power	LPM	Listen power	T.power	Power	LDC	TDC
1.0	1.000	8.000	1.000	0.387	0.154	0.485	0.109	1.132	0.805	0.200
2.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3.0	1.000	8.272	1.035	0.392	0.484	0.481	0.083	1.104	0.815	0.143
4.0	1.001	18.906	2.362	0.367	0.406	0.409	0.093	1.016	0.620	0.175
5.0	2.000	61.001	0.910	0.378	0.452	0.452	0.101	1.081	0.705	0.188
6.0	2.454	30.544	2.366	0.372	0.416	0.416	0.094	1.035	0.625	0.179
7.0	3.000	45.816	2.365	0.375	0.430	0.430	0.095	1.050	0.705	0.172
8.0	2.002	28.637	2.581	0.378	0.411	0.412	0.081	1.022	0.655	0.152
9.0	2.000	61.005	0.910	0.364	0.409	0.412	0.092	1.018	0.618	0.172
10.0	1.000	18.910	2.366	0.379	0.438	0.438	0.068	1.034	0.711	0.120
Average	1.546	28.109	1.590	0.339	0.360	0.394	0.082	0.949	0.626	0.150

The simulation results in details have shown in Tables 6, 7, 8, and 9 corresponding random, linear, ellipse, and manual positioning system of motes. The simulations have taken 10, 20, and 30 motes, when the numbers of motes increase to communicate, then more power consumption is required than a small number of motes share in the networks.

All these practical information based on RPL and 6LoWPAN protocols. This detailed analysis of RPL and 6LoWPAN protocols will help explore more options to develop new protocols and enhance these existing protocols like mobile motes in real-time development of IoT networks. Also, we simulate the case of static motes with various types of real motes and various network topology by Eq. 1 and 2.

$$Power\ Consumption\& = \frac{Energest_Value * Current * Voltage}{RTIMER_SECOND * Runtime} \quad (1)$$

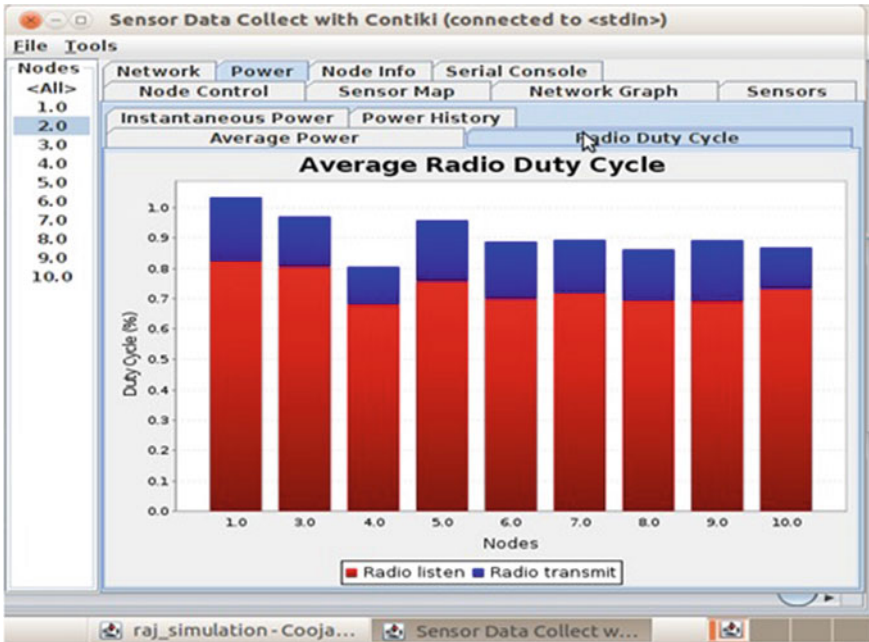


Fig. 7 Average radio duty cycle

$$Duty\ Cycle(\%) \& = \frac{Energest_TX + Energest_RX}{Energest_CPU + Energest_LPM} \tag{2}$$

So, all the program used in Contiki 2.7 with “powertrace.h” files and “make files”. Include powertrace application in our application by adding the Makefile.
 APP+ = powertrace
 # include “powertrace.h”
 Add to source file to print power profile every 5 s:
 powertrace_start(CLOCK_SECOND*5)

4.6 Observation of Mote’s Lifetime

The results comparing motes in the Internet of things with various types of topology configurations show the graph from Figs. 5 and 6 from the experiments. We have obtained that motes times with different values, which guide in Table 6. Our analysis is the motes duty cycle, power consumption, CPU power, EXT of each mote. In the IoT system, we will try to find out the failure of motes power, duty cycle, transmission, etc., and calculate the power and duty cycle percentage from Eqs. 1 and 2. Table 6 shows all types of values of each node or notes information. Figure 12 shows are the

mote's CPU power, motes power consumption, the low duty cycle of mote's and EXT concerning motes. And other values you can calculate by tables as your requirements; in our application, we calculate the CPU power of motes in like random positioning (RP), linear positioning (LP), ellipse positioning (EP), and manual positioning (MP). The motes how much consume power. The data shown in the same tables with various type topology, expected transmission count (EXT) form the same tables, and low duty cycle calculated in IoT. In any topology construction, not more value changes could calculate values from questions 1,2 and other things of motes calculated by Tables 6, 7, 8, and 9.

5 Research Challenges of IoT

IoT brings various challenges ranging after the simulation. We find out different research challenges which come under to customization and some improvement of results after implementation. These beneficial in home automation, industrial, security, dynamic changing the network, power, reduce duty cycle after changing the protocol services, address compression, dynamic address changes according to devices to devices [28, 30]. Like device integration, heterogeneity, scalability to mobility, routing, security, and other specific challenges. It is clear from the above that IoT security is an aspect that requires in-depth research work as the need to secure networks today has become imperative [22, 23].

6 Conclusions

In this work, we have measured the performance of motes in different scenarios of topology constructions and for many matrices. The experimental results illustrate the more values of sky motes. Furthermore, the motes' performance, not more, is changing when they are any topology in networks. In this paper, comparisons have made using different matrices: number of hops, low power mode (LPM), EXT, and many types of power consumption of motes like CPU power, average power, listen to authenticity, duty cycle power. The simulations have 10, 20, and 30 motes; when the number of motes increases to communicate, more power consumption requires than a small number of motes share in the networks. These detailed analysis of motes based on RPL and 6LoWPAN protocols will help explore more options to develop new protocols and enhance these existing protocols like mobile motes in real-time development of IoT networks. Also, we will investigate the case of static motes with various types of the entire network. All the programs used in Contiki 2.7 with power trace.h files and make files. Include power trace application in our application by adding the Makefile.

Observation of mote's lifetime: The results comparing motes in the Internet of things with various types of topology configurations and the graph Figs.5 and 6 from the experiments. We have obtained that motes times with different values, which guide in Table 6. Our analysis is base on the motes duty cycle, power con-

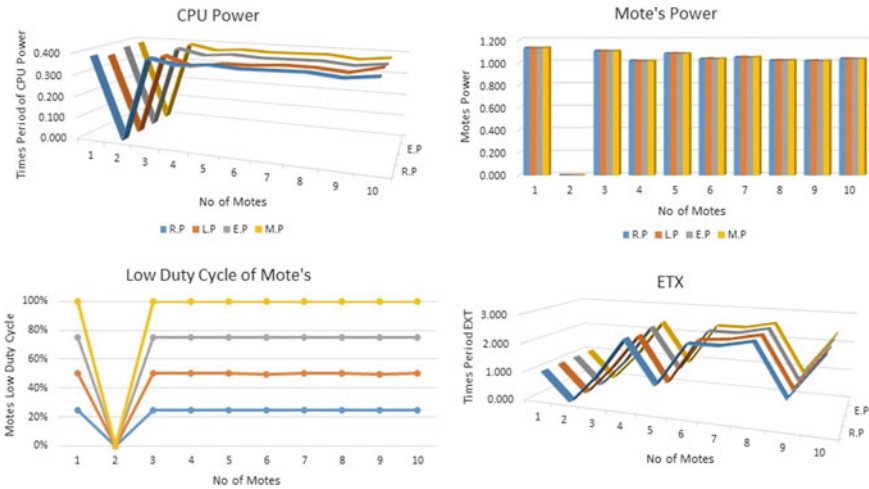


Fig. 8 Relation between CPU power, motes power, EXT, and duty cycle with different topology constructions

sumption, CPU power, and EXT. In the IoT system, we will determine the failure of motes power, duty cycle, transmission, and calculate the power and duty cycle percentage from Eqs. 1 and 2. Table 6 shows the various values of each mote when they communicate with others. Figure 8 shows the mote's CPU power, motes power consumption, the low duty cycle of mote's, and EXT concerning motes. Moreover, other values can calculate by tables as various topology; in our application, we calculate the values in many positing as random positioning (RP), linear positioning (LP), ellipse positioning (EP), and manual positioning (MP). The motes data show in the tables with various topologies and different values calculate in IoT applications. So, the results of our simulations show in Tables 6, 7, 8, and 9.

Abbrevation

The following abbreviations are used in this manuscript:

LP	Listen power
LPM	Low power mode
TP	Transmission power
AIPT	Avg. inter-packet time (min-sec)
MIPT	Min. inter-packet time (sec)
MAIPT	Max. inter-packet time
LDC	Listen duty cycle
TDC	Transmission duty cycle
LP	Linear positioning
RP	Random positioning
EP	Ellipse positioning
MP	Manual positioning
EXT	Expected transmission count

References

1. D. Airehrour, J.A. Gutierrez, S.K. Ray, Sectrust-rpl: a secure trust-aware rpl routing protocol for Internet of Things. *Future Generation Comput. Syst.* **93**, 860–876 (2019)
2. S. Chakraborty, A. Majumder, 6lowpan security: classification, analysis and open research issues. *Int. J. Comput. Intell. IoT* **1**(1) (2019)
3. M. Cooperation, Tmote sky: ultra low power ieee 802.15.4 compliant wireless sensor module. Datasheet: <http://www.sentilla.com/pdf/eol/tmote-skydatasheet.pdf> (2005)
4. D.W. Courtney, P. Thulasiraman, Implementation of secure 6lowpan communications for tactical wireless sensor networks, in *2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 704–709, IEEE (2016)
5. D. Djenouri, I. Balasingham, Power-aware qos geographical routing for wireless sensor networks—implementation using contiki, in *2010 6th IEEE International Conference on Distributed Computing in Sensor Systems Workshops (DCOSSW)*, pp. 1–5, IEEE (2010)
6. A. Dunkels, Rime—a lightweight layered communication stack for sensor networks, in *Proceedings of the European Conference on Wireless Sensor Networks (EWSN), Poster/Demo session*, Delft, The Netherlands, vol. 44. Citeseer (2007)
7. A. Dunkels, J. Eriksson, N. Finne, N. Tsiftes, *Powertrace: Network-Level Power Profiling for Low-Power Wireless Networks* (2011)
8. A. Dunkels, B. Gronvall, T. Voigt, Contiki—a lightweight and flexible operating system for tiny networked sensors, in *29th Annual IEEE International Conference on Local Computer Networks*, pp. 455–462, IEEE (2004)
9. A. Dunkels, F. Österlind, Z. He, An adaptive communication architecture for wireless sensor networks, in *Proceedings of the 5th International Conference on Embedded Networked Sensor Systems*, pp. 335–349 (2007)
10. A. Dunkels, O. Schmidt, T. Voigt, M. Ali, Protothreads: simplifying event-driven programming of memory-constrained embedded systems, in *Proceedings of the 4th International Conference on Embedded Networked Sensor Systems*, pp. 29–42 (2006)
11. J. Eriksson, F. Österlind, N. Finne, A. Dunkels, T. Voigt, Accurate power profiling for sensor network simulators, in *8th Scandinavian Workshop on Wireless Ad-hoc & Sensor Networks*, Stockholm, Sweden, May 2008
12. J. Eriksson, F. Österlind, N. Finne, N. Tsiftes, A. Dunkels, T. Voigt, R. Sauter, P.J. Marrón, Cooja/mspsim: interoperability testing for wireless sensor networks, in *Proceedings of the 2nd International Conference on Simulation Tools and Techniques*, pp. 1–7 (2009)
13. M.M. Hossain, S., Al Noor, Y. Karim, R. Hasan, Iotbed: a generic architecture for testbed as a service for internet of things-based systems, in *ICIOT*, pp. 42–49 (2017)
14. L. Kanaris, C. Sergiou, A. Kokkinis, A. Pafitis, N. Antoniou, S. Stavrou, On the realistic radio and network planning of IoT sensor networks. *Sensors* **19**(15), 3264 (2019)
15. K.F. Krentz, H. Rafiee, C. Meinel, 6lowpan security: adding compromise resilience to the 802.15.4 security sublayer, in *Proceedings of the International Workshop on Adaptive Security*, pp. 1–10 (2013)
16. R. Lajara, J. Pelegrí-Sebastiá, J.J.P. Solano, Power consumption analysis of operating systems for wireless sensor networks. *Sensors* **10**(6), 5809–5826 (2010)
17. S.K. Lee, M. Bae, H. Kim, Future of IoT networks: a survey. *Appl. Sci.* **7**(10), 1072 (2017)
18. N. Mishra, L.P. Verma, P.K. Srivastava, A. Gupta, An analysis of IoT congestion control policies. *Proc. Comput. Sci.* **132**, 444–450 (2018)
19. F.D. Miyandoab, J.C. Ferreira, V.M.G. Tavares, Analysis and evaluation of an energy-efficient routing protocol for wsns combining source routing and minimum cost forwarding. *J. Mobile Multimedia* **14**(4), 469–504 (2018)
20. F. Osterlind, A. Dunkels, J. Eriksson, N. Finne, T. Voigt, Cross-level sensor network simulation with cooja, in *Proceedings. 2006 31st IEEE Conference on Local Computer Networks*, pp. 641–648, IEEE (2006)
21. A. Patil, R.V. Biradar, Programming the Sensor Nodes in WSN

22. S. Prakash, A. Rajput, Hybrid cryptography for secure data communication in wireless sensor networks, in *Ambient Communications and Computer Systems*, pp. 589–599, Springer (2018)
23. S. Prakash, V. Saroj, A review of wireless charging nodes in wireless sensor networks, in *Data Science and Big Data Analytics*, pp. 177–188, Springer (2019)
24. T. Preiss, M. Sherburne, R. Marchany, J. Tront, Implementing dynamic address changes in contikiOS, in *International Conference on Information Society (i-Society 2014)*, pp. 222–227, IEEE (2014)
25. D. Shreenivas, S. Raza, T. Voigt, Intrusion detection in the rpl-connected 6lowpan networks, in *Proceedings of the 3rd ACM International Workshop on IoT Privacy, Trust, and Security*, pp. 31–38 (2017)
26. E. Siow, T. Tiropanis, W. Hall, Analytics for the internet of things: a survey. *ACM Comput. Surv. (CSUR)* **51**(4), 1–36 (2018)
27. S.S. Solapure, H.H. Kenchannavar, Rpl and coap protocols, experimental analysis for iot: a case study. *Int. J. Ad hoc* **10**(2) (2019)
28. S. Srivastava, S. Prakash, Security enhancement of IoT based smart home using hybrid technique, in *International Conference on Machine Learning, Image Processing, Network Security and Data Sciences*, pp. 543–558, Springer (2020)
29. A. Velinov, A. Mileva, Power consumption analysis of application layer protocols for the Internet of Things, in *International Conference on ICT Innovations*, pp. 193–202, Springer (2016)
30. G. Verma, S. Prakash, Internet of Things for healthcare: research challenges and future prospects, in *Advances in Communication and Computational Technology*, pp. 1055–1067, Springer
31. O. Vermesan, P. Friess, *Internet of Things: Converging Technologies for Smart Environments and Integrated Ecosystems*. River publishers (2013)
32. T. Voigt, F. Osterlind, A. Dunkels, *Contiki cooja hands-on crash course: Session notes*. Swedish Institute of Computer Science, pp. 1–10 (2009)
33. A. Yeole, D. Kalbande, A. Sharma, Security of 6lowpan IoT networks in hospitals for medical data exchange. *Proc. Comput. Sci.* **152**, 212–221 (2019)
34. Y.B. Zikria, S.W. Kim, O. Hahm, M.K. Afzal, M.Y. Aalsalem, *Internet of Things (IoT) Operating Systems Management: Opportunities, Challenges, and Solution* (2019)

Fruit Classification Using Deep Learning



Jinu Lilly Joseph, Veena A. Kumar, and Santhosh P. Mathew

Abstract Fruit classification plays an important role in many industrial applications including factories, supermarkets and other fields. The importance of fruit classification can also be seen among people with dietary requirements to assist them in selecting the correct categories of fruits. Manual sorting method for fruit classification was used earlier which is time consuming and involves human presence always. Many machine learning approaches for fruit classification were proposed in the past. Deep learning can be a powerful engine for producing actionable results in today's world with its abilities of recognition and classification. Thus, an efficient model for fruit classification was developed using convolutional neural network in deep learning. It uses fruits 360 dataset containing 131 different classes of fruits and vegetables that can classify images into different categories. The model was developed using TensorFlow backend. It was trained using 50 epochs, and it obtained an accuracy of 94.35%.

Keywords Fruit classification · TensorFlow · Convolutional neural network · Deep learning

1 Introduction

In artificial intelligence, deep learning and machine learning are closely related. In other words, deep learning can be viewed as a subset of machine learning. It is based on artificial neural networks, and it tries to mimic the human brain in various activities like pattern recognition, image classification, speech recognition, game playing, text classification and so on. We can see tremendous potential of deep learning in many areas like education, medicine, business and so on. In deep learning, the system tries to learn from examples rather than being explicitly programmed.

Image classification can be seen as an important application of deep learning. It becomes necessary to classify images in many applications. One among them is

J. L. Joseph (✉) · V. A. Kumar · S. P. Mathew
Saintgits College of Engineering, Kottukulam Hills, Pathamuttom, Kottayam, India
e-mail: jinu.css1921@saintgits.org

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_62

807

fruit classification. Automatic fruit classification has its importance, especially in industries. It is used to help the cashiers in supermarkets to identify various fruits with the intention of minimal wastage of time and energy. It is also important for people with dietary requirements and health issues to identify various species of fruits that suit their health requirements.

We propose a deep learning model using convolutional neural network for classifying various images of fruits accurately. The proposed model for performing classification is defined based on the sequential model of keras, and it consists of a set of five convolutional layers, five pooling layers, flattened matrix, fully connected layers and finally a softmax layer. The model was trained using 50 epochs on Fruits 360 dataset available on kaggle, and it classifies the images of fruits into 131 different categories. The model obtained an accuracy of 94.35%.

The rest of the paper is arranged as follows: Sect. 2 explains the related works in fruit classification from the literature; Sect. 3 contains the problem definition; Sect. 4 explains our approach along with the results; Sect. 5 contains the conclusions.

2 Literature Overview

In the work done by M. S. Hossain and his colleagues [1], the importance of fruit classification in today's world is depicted by proposing efficient frameworks using two different datasets. Two different models are proposed for the same. The first model is a proposed light model and the second model is a pre-trained visual geometry group-16 model. Using dataset 1, the accuracies obtained were 99.49% and 99.75%, and using dataset 2, the accuracies were 85.43% and 96.75% for the two models. It also describes the implementation of the fruit classification system as a mobile application for the ease of use, especially among customers with dietary requirements in identifying the necessary fruit species. The application has the ability of identifying the different fruit species along with the necessary details.

In the work done by Sakib S and the other authors [2], the recognition of fruits from images using convolutional neural network is depicted. The model used consists of two convolutional layers. They are followed by the pooling layers and fully connected layers. 64 kernels were used, and the rectified linear unit was the activation function used. The images of Fruits 360 dataset were used in the model.

Y.-D. Zhang and his colleagues [3] proposed a 13-layer convolutional neural network for performing the classification of fruits. The fruit dataset was developed through on-site collecting using digital camera, downloading the images from Google and so on. Finally, a dataset containing 3600 images was obtained. The convolutional layer performs the convolution operation between input images and filters. Using the data augmentation on the image dataset, the training data was expanded from 1800 to 63,000. The proposed convolutional neural network (CNN) model was able to obtain an accuracy of 94.94%.

In the work done by S.R. Dubey and the other authors [4], a framework for recognizing fruits and vegetables is proposed. The model takes the images of fruits

and vegetables as input, and the species and variety will be given as the output. The problem of classifying fruits and vegetables can be seen in mostly among the supermarkets. Thus, this problem is solved by identifying various fruits purchased by the customers easily.

In the work done by Zhang Y. and the other authors [5], a multi-class kernel support vector machine (kSVM) is incorporated for the purpose of classifying the fruits with very good accuracy. The image segmentation is performed before the actual classification. The color distribution in the image is performed with the help of color histogram. The dimensionality reduction is performed using the principal component analysis technique. The results show that the Max-Wins-Voting Support Vector Machine (MWV-SVM) with Gaussian Radial Basis(GRB) kernel obtained the best classification accuracy of 88.2%.

In the work done by A. Rocha and the other authors [6], a unified approach which combined many features and classifiers is used. The classification problem containing multiple classes is approached as a set of binary problems. This method also requires less training.

3 Problem Definition

The importance of fruit classification can be seen in many fields such as industries and also among people with dietary requirements and health issues [7–9]. The traditional method of fruit classification which is the manual sorting is a daunting task due to the chances for errors and reduced human performance. In order to avoid these drawbacks of manual sorting and machine learning methods, a deep learning approach is proposed in this paper for automatic fruit classification [10, 11].

The Fruits 360 dataset containing high-quality images of fruits and vegetables is the dataset used for both the training and the testing purpose. The convolutional neural network which belongs to the class of deep neural networks is used based on the sequential model of keras and TensorFlow [12, 13].

4 Proposed Work

Hereinafter, we describe the dataset in Sect. 4.1. Section 4.2 describes the details of dataset augmentation. Section 4.3 explains the proposed convolutional neural network architecture. Section 4.4 deals with the tools and frameworks employed in this classification. Section 4.5 includes the results obtained from experiments and discussions.

The main steps of this fruit classification system are:

1. Data loading,
2. Model creation,

3. Training of the model,
4. Testing of the model.

The dataset used was Fruits 360 dataset available on kaggle. It is organized as training and test folders of high-quality images of fruits and vegetables. The images were subjected to data augmentation during the pre-processing step to avoid the problem of overfitting.

The proposed model was defined based on sequential model of keras. The model was composed of a set of five convolutional layers. A pooling layer follows every convolutional layer. The convolution operation is performed between each input fruit image and the filters to extract various features from the fruit image. After every convolution layer, the number of filters was increased with the purpose of extracting more and more features from the input image.

The pooling operation is done after each convolution layer in order to reduce the size of the image which in turn reduces the number of nodes for flattening and to the fully connected network. The activation function used in the model was the rectified linear unit in order to eliminate the negative values in the feature vector and to retain only the positive values. This is followed by flattening of the resulting matrix into a single vector, and it is passed to the fully connected layers. In addition to the flattened input matrix, the fully connected layers also have the hidden layers and output layers.

Finally, a softmax layer containing a softmax function is added which computes the probabilities of various classes and assigns each input image to the class with the highest probability.

Then, the convolutional neural network model was fitted to the dataset for enabling it to fit to the training and test image datasets. The model was trained using 50 epochs, and it achieved a training accuracy of 99.32%, and the testing was performed, and the test accuracy was found to be 94.35%.

4.1 Fruits 360 Dataset

The Fruits 360 dataset containing high-quality images of fruits organized into training and testing folders is used as the dataset for the model. This is again sub-divided into various folders of fruit classes. This structure is required for importing images using keras.

The images from the dataset are subjected to data augmentation during the pre-processing stage in order to avoid the problem of overfitting. The dataset contains a total of 90,483 images of fruits and vegetables. It includes 131 categories of different fruits and vegetables (Fig. 1).

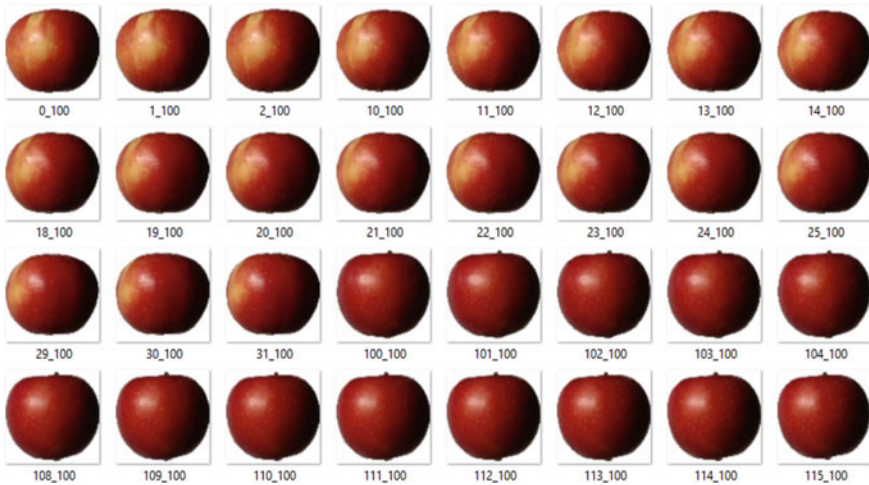


Fig. 1 Sample of Fruits 360 dataset

4.2 Dataset Augmentation

In order to reduce the problem of overfitting where a model works well on the training set than on the test set, dataset augmentation is performed which applies various transformations on the existing Fruits 360 training dataset. The transformations include flipping, rotating, zooming, shearing, rescaling and so on. The ImageDataGenerator function of keras is used to perform the dataset augmentation.

4.3 Convolutional Neural Network (CNN) Architecture

The deep learning approach for fruit classification was performed using the sequential model of keras to initialize the convolutional neural network. The convolutional neural network for fruit classification consists of five convolutional layers. Every convolutional layer is followed by a pooling layer which performs max pooling (Figs. 2 and 3).

Each convolutional layer performs the convolution operation between the input fruit image and the filters resulting in the production of feature maps. The number of kernels or filters was increased after each convolution layer with the purpose of extracting more and more features from the images. The number of kernels of the five convolutional layers is in the order of 16, 32, 48, 64 and 128.

The activation function used was the rectifier linear unit (ReLU) which removes all the negative values from the network since negative values do not contribute much to the network. ReLU activation function is used in every convolutional layer.

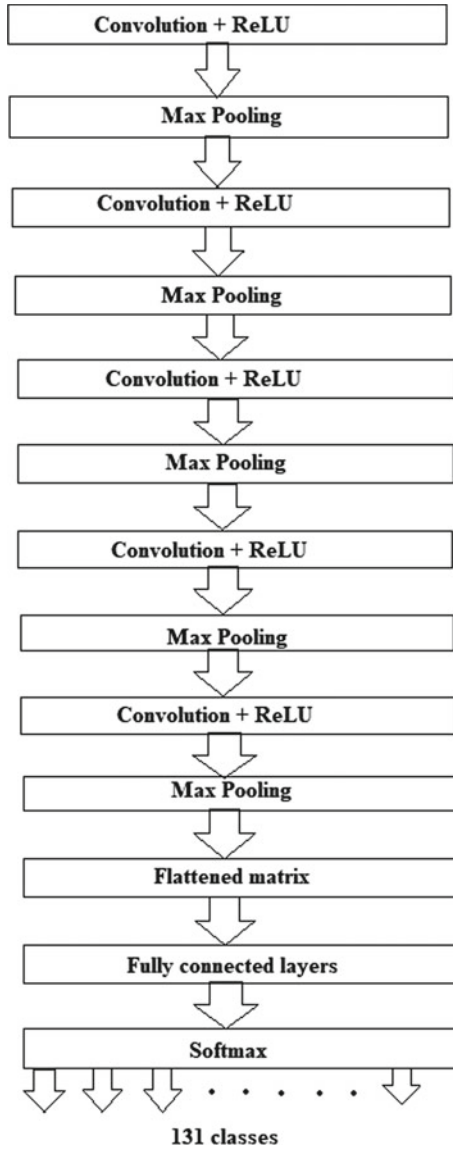


Fig. 2 Proposed CNN architecture

For an input value x , the ReLU activation function can be defined mathematically as:

$$f(x) = \max(0, x).$$

[39] Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	multiple	448
max_pooling2d (MaxPooling2D)	multiple	0
conv2d_1 (Conv2D)	multiple	4640
max_pooling2d_1 (MaxPooling2D)	multiple	0
conv2d_2 (Conv2D)	multiple	13872
max_pooling2d_2 (MaxPooling2D)	multiple	0
conv2d_3 (Conv2D)	multiple	27712
max_pooling2d_3 (MaxPooling2D)	multiple	0
conv2d_4 (Conv2D)	multiple	73856
max_pooling2d_4 (MaxPooling2D)	multiple	0
flatten (Flatten)	multiple	0
dense (Dense)	multiple	67203

=====
Total params: 187,731
Trainable params: 187,731
Non-trainable params: 0

Fig. 3 Summary of model

There is a pooling layer after each convolution layer. The max pooling is used which chooses the maximum value from the window in the feature maps. Thus, pooling helps in reducing the size of the feature maps produced as a result of the convolution of input fruit images and kernels. The stride is defined as the number of pixels shifted over the input matrix. The stride for pooling operation is 2.

After the fifth pooling layer, the resulting matrix was flattened. Then, it was passed through the fully connected layers. Finally, a softmax layer was used which assigns each of the given image to the corresponding class. The softmax layer contains a softmax function which computes the probabilities of each class for a given image and assigns the image to the class with the highest probability.

Then, the model was compiled using adam optimizer which is a stochastic gradient algorithm, and the loss was given as categorical cross-entropy since the classification problem comprised of more than two categories. The performance metric was assigned with accuracy (Table 1).

Table 1 Hyperparameters of the model

Loss function	Categorical cross-entropy
Batch size	128
Epochs	50
Optimizer	Adam
Activation functions	ReLU and softmax

4.4 Tools and Frameworks

Google colab

Google colab (colab) is a Jupyter-based notebook which helps in running machine learning and deep learning projects, and it runs on cloud. It provides support for graphics processing unit (GPU) in addition to tensor processing units (TPU) and central processing unit (CPU). It is free and easy to use since it avoids the complex installations unlike other tools.

Python

Python language which is a simple and common language is used. It is simple and easy to understand, and hence, it is used for building models easily in deep learning. It does not have any confusing syntax, and hence, it saves a lot of time in project development.

TensorFlow

It is a library of symbolic computation developed by Google, and it is used in many machine learning applications. It is open source, and it provides a platform for creating various machine learning models as well as deep learning models. It is a library for fast numerical computing, and it is easy to understand.

Keras

It is a free and open-source library written in Python, and it provides models like sequential model and so on. It is an API, and it is possible to import images with keras.

It runs on the top of TensorFlow. It is a popular and commonly used deep learning framework.

4.5 Experimental Results and Discussions

Deep learning has tremendous potential in almost all areas. The fruit classification system using convolutional neural network can be adapted to suit many applications which require the identification of specific fruits.

The proposed using convolutional neural network (CNN) model is composed of five convolutional layers. A pooling layer follows every convolutional layer. The ReLU was used as the activation function in each convolution layer. Following the convolutional layers and the pooling layers, the model also has flattened matrix, fully connected layers and a softmax layer. The model was trained using 50 epochs on the Fruits 360 dataset. The model obtained an accuracy of 99.32% on training dataset and 94.35% on testing (Table 2; Figs. 4 and 5).

Table 2 Comparison table based on the experiments conducted by Muresan et al. [14]

Configuration	Accuracy (%)
Convolutional neural network with four convolution layers (16, 32, 32 and 128 kernels) and two fully connected layers (1024 and 256 nodes)	92.30
Convolutional neural network with four convolution layers (16, 16, 64 and 128 kernels) and two fully connected layers (1024 and 256 nodes)	93.13
Convolutional neural network with four convolution layers (16, 32, 64 and 128 kernels) and two fully connected layers (512 and 256 nodes)	93.26
Convolutional neural network with four convolution layers (16, 32, 64 and 128 kernels) and two fully connected layers (1024 and 512 nodes)	94.16
Proposed system	94.35

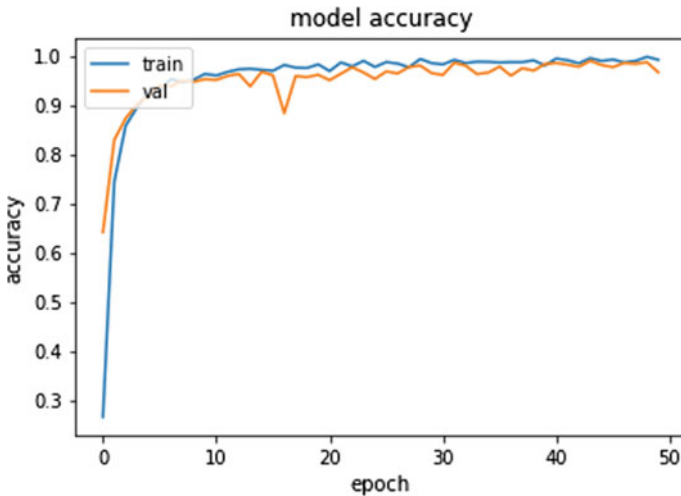


Fig. 4 Curve of model accuracy

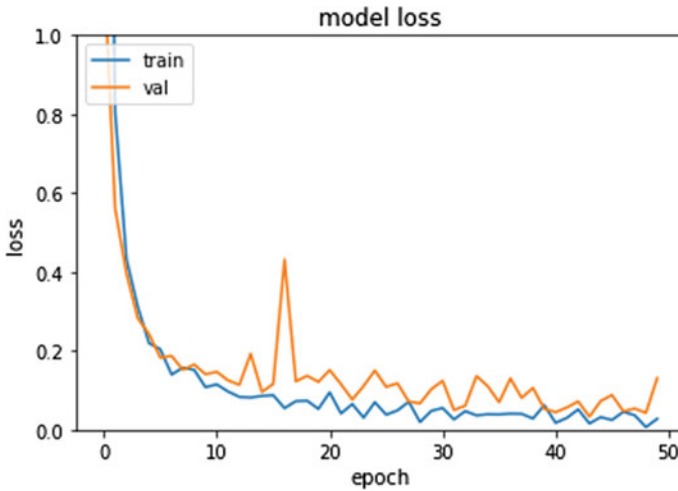


Fig. 5 Curve of model loss

5 Conclusions

The deep learning approach for fruit classification is suitable for many useful applications like industry, education and so on. The model was developed using convolutional neural network in deep learning. The model was trained using 50 epochs on Fruits 360 dataset, and it obtained an accuracy of 94.35%. The tremendous impact of deep learning in today's world and its ability to process more and more features help in solving the drawbacks of the traditional methods.

References

1. M.S. Hossain, M. Al-Hammadi, G. Muhammad, Automatic fruit classification using deep learning for industrial applications. *IEEE Trans. Ind. Informat.* **15**(2), 1027–1034 (2019)
2. S. Sakib, Z. Ashrafi, M.A. Siddique, *Implementation of Fruits Recognition Classifier using Convolutional Neural Network Algorithm for Observation of Accuracies for Various Hidden Layers* (2019)
3. Y.-D. Zhang et al., Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. *Multimedia Tools Appl.*, 1–20, Sep. 2017
4. S.R. Dubey, A.S. Jalal, Fruit and vegetable recognition by fusing colour and texture features of the image using machine learning. *Int. J. Appl. Pattern Recognit.* **2**(2), 160–181 (2015)
5. Y. Zhang, L. Wu, Classification of fruits using computer vision and a multiclass support vector machine. *Sensors* **12**(9), 12489–12505 (2012)
6. D.C. Rocha, J. Hauagge, S. Wainer, Goldenstein, Automatic fruit and vegetable classification from images. *Comput. Electron. Agric.* **70**(1), 96–104 (2010)
7. S.R. Dubey, A.S. Jalal, Robust approach for fruit and vegetable classification. *Proc. Eng.* **38**, 3449–3453 (2012)

8. S. Jana, R. Parekh, Shape-based Fruit Recognition and Classification, in *International Conference on Computational Intelligence, Communications, and Business Analytics*, Singapore, Mar. 24 2017
9. K. Kheiralipour, A. Pormah, Introducing new shape features for classification of cucumber fruit based on image processing technique and artificial neural networks. *J. Food Process Eng.* e12558, Mar. 1, 2017. <https://doi.org/10.1111/jfpe.12558>
10. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering* ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh. Lecture Notes in Electrical Engineering, vol 661 (Springer, Singapore) (2021). https://doi.org/10.1007/978-981-15-4692-1_58
11. G. Karpathy, S. Toderici, T. Shetty, R. Leung, Sukthankar, L. Fei-Fei, Large-scale video classification with convolutional neural networks, in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 1725–1732 (2014)
12. S. Mandal, S. Biswas, V.E. Balas, R.N. Shaw, A. Ghosh, Motion prediction for autonomous vehicles from lyft dataset using deep learning, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, 30–31 Oct. 2020, pp. 768–773. <https://doi.org/10.1109/ICCCA49541.2020.9250790>
13. H. Cheng, L. Damerow, Y. Sun, M. Blanke, Early yield prediction using image analysis of apple fruit and tree canopy features with neural networks. *J. Imaging* **3**, 6 (2017)
14. H. Mureşan, M. Oltean, Fruit recognition from images using deep learning. *Acta Universitatis Sapientiae. Informatica* **10**(1), 26–42

Orthonormal Bayesian Convolutional Neural Network for Detection of the Novel Coronavirus-19



Romin Chetankumar Jariwala and Madhusudana Rao Nalluri

Abstract Novel coronavirus-19 (COVID-19) is almost affected all over the world, and daily, thousands of peoples are getting infected and died due to this virus. All the existing methods to test coronavirus infections are taking a long time to give results and giving low sensitivity and specificity. This necessitates artificial intelligence (AI) in the detection of COVID-19 infected patients. Various deep learning and image processing techniques are available to recognize patterns from the X-ray and the computed tomography (CT) image and giving higher testing accuracy, sensitivity, and specificity on a large dataset. In this research, we adapted the Bayesian inference approach and proposed a Bayesian convolutional neural network (BCNN) with an orthogonal normalization technique (ONT) for automatic detection of the COVID-19, pneumonia, and no-infection (Normal) from CT images. To verify the strength of the model, the data is randomly split up into different training, validation, and testing set combinations. With the proposed model on (training%_validation%_testing%) 70_10_20 split up, we achieved 99.91% testing accuracy with 99.91% F1 score and 100% area under the receiver operating characteristic curve (AUC). We also achieved 100% sensitivity, 100% specificity on COVID-19 prediction, 100% sensitivity, 99.97% specificity on normal prediction, and 100% sensitivity, 99.91% specificity on the pneumonia prediction. The higher accuracy, sensitivity, and specificity of this model would be extremely useful for the accurate and rapid screening of the COVID-19 and pneumonia.

Keywords COVID-19 · Pneumonia · Bayesian inference · CNN · Orthogonal normalization

R. C. Jariwala · M. R. Nalluri (✉)
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: n_madhusudanarao@cb.amrita.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_63

819

1 Introduction

COVID-19 was first identified in China's Wuhan City, in December 2019, and it was caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1]. This outbreak was declared a pandemic by the World Health Organization (WHO) on March 11, 2020 [2]. As of October 18, 2020, this virus has spread to 215 countries with a total of 40,023,720 cases, 1,115,601 total deaths, and 29,935,273 recovered cases [3]. If we check statistics for each country, the USA is on the top with 8,343,244 cases, and India is in the second place with 7,494,551 cases [3]. Generally, this virus is spread through contact and respiratory droplets, and less immune people are most affected by this virus. Since no vaccine is available to date October 18, 2020, it is very difficult to deal with this situation.

Currently, four medical tests are available to detect the SARS-CoV-2 virus. The frequently used test is the reverse transcription—polymerase chain reaction (RT-PCR) test [4], which uses nasal and throat swabs to detect the presence of the virus in the human body. RT-PCR test is used for early detection of the SARS-CoV-2 virus since it is detecting the ribonucleic acid (RNA) of the virus, and this test takes 4 to 5 h to get results with 60–80% sensitivity and 90–95% specificity. RT-PCR test is more accurate, and we can test 90 samples at the same time [4]. Another test is TrueNat and cartridge-based nucleic acid amplification test (CBNAAT), its working principle is the same as the RT-PCR test, and it is giving results in 60 min, with 50–80% sensitivity and 90–95% specificity. Only one to four samples can be tested at a time with this method [5]. To detect antigen, new method was approved by the Indian Council of Medical Research (ICMR), called the rapid point-of-care (PoC) antigen detection test which is useful for diagnosis with RT-PCR test. This test is taking 15 to 30 min to give a result with, 50.6–84% sensitivity and 93.3–100% specificity [6]. For surveillance purposes, the immunoglobulin G (IgG) antibody test is used. This method is used to find whether previously this individual is infected with SARS-CoV-2 or not. This method is taking 20 to 30 min to produce results with 92.7% sensitivity and 97.9% specificity [7]. All these four techniques require a long process, less accurate, and takes a long time to generate results. The COVID-19 infection is also detected by computed tomography (CT) scans and X-ray images to increase sensitivity and reduce time. According to the study [8], a chest CT scan is more sensitive than a radiograph. Because of the more sensitivity, CT scan images are more likely to use detect COVID-19 infection. Since the COVID-19 disease is one type of pneumonia, it is a challenging task to differentiate between pneumonia and COVID-19. Mostly, ground-glass opacification pattern (GGO) is used to find COVID-19 infections in CT images which is most commonly located in the inferior lobe of the right lung, and it is presented as a unifocal lesion in the early stage of the disease [9].

In this paper, we proposed the deep learning method to classify COVID-19, pneumonia, and normal using CT images. COVID-19 is another type of pneumonia only so, efficient classification of COVID-19 and pneumonia is very helpful to speedup tests and enhances the sensitivity of detection. To preprocess images, we introduce

a new pipeline that contains an image enhancement technique based on histogram equalization and gamma transform. Furthermore, we also applied image augmentation techniques to introduce more randomization in training data. For the classification of the images, we proposed a Bayesian convolutional neural network (BCNN) with an orthogonal normalization technique (ONT) and show how ONT is increasing the performance of the network and giving higher testing accuracy, sensitivity, specificity, F1 score, and receiver operating characteristic curve (AUC). Using an orthogonal weight matrix, we normalized the learned weights to optimize and improve the performance of the model [10]. For faster learning of the network, the RMSprop optimizer is used to train the model. The proposed approach does not only give better performance than all existing methods [4–7], but it is also converging faster. So, it is very helpful for a quick diagnosis of the COVID-19 and pneumonia.

Main technical contributions of this paper are

1. The preprocessing pipeline, which contains image enhancement and image augmentation techniques to efficiently detect COVID-19 from CT images.
2. Modified AlexNet architecture with Bayesian orthogonal normalized layers.

The rest of the paper is structured as follows: The work in literature is explained in Sect. 2, The proposed methodology is explained in Sect. 3, which contains an in-depth explanation of the preprocessing pipeline and proposed modified AlexNet architecture with Bayesian orthogonal normalized layer, The dataset information and the entire details of the training process are explained in Sect. 4. Best outcome and some experimental results are explained in Sect. 5, and in Sect. 6, we conclude this work.

2 Literature Survey

To diagnose lung disease efficiently in less amount of time, it is a challenging task in medical science. Sometimes visual examination of the CT or X-ray images leads to an inefficient result. To overcome this issue, automated disease detection techniques are developed with the latest advancement in medical image processing and innovative deep learning architectural designs. Many new techniques come up to diagnose lung diseases such as Xiaosong et al. [11] proposed a method to classify thorax diseases using X-ray images, Stephen et al. [12] proposed a method to detect emphysema pattern in lungs, Carolina et al. [13] proposed a method for asthma detection using CT images, Lisa et al. [14] proposed an approach to detect chronic obstructive pulmonary disease (COPD) in lungs using low dose CT images, Qin et al. [15] proposed tuberculosis detection method, and Senthil et al. [16] proposed various evolutionary algorithms to detect lung cancer. To diagnose disease other than lungs, Poorna et al. [17] proposed melanoma detection using deep learning, Saiprasath et al. [18] compare malaria detection algorithms using microscopic images, Kiruthika et al.

[19] proposed the artery and vein classification for retinopathy and to detect microaneurysm in fundus images, and Kumar et al. [20] proposed the feature selection techniques using particle swarm optimization.

COVID-19 and pneumonia classification is a challenging task because in the initial phase the COVID-19 symptoms are the same as pneumonia. So, if we can detect pneumonia in early stage and scan out those infected people, then it will be very helpful to overcome the spread of the virus. To detect pneumonia in the early stage, many kinds of research come out. Hashmi et al. [21] proposed a novel approach to combine weighted predictions from the ResNet18, Xception, InceptionV3, DensNet121, and MobileNetV3 models. They achieved 98.43% test accuracy and 99.76 AUC score. A transfer learning-based approach for pneumonia detection is proposed by Vikash et al. [22]. They applied transfer learning using AlexNet, DensNet121, InceptionV3, ResNet18, and GoogLeNet neural networks. On the ensemble model, they achieved 96.4% accuracy and 99.62% recall. Without preparation of the group, Garima et al. [23] proposed a CNN model to classify pneumonia. They achieved 92.89% testing accuracy. Praveen et al. [24] used an unsupervised fuzzy c-means algorithm for the classification of pneumonia, and they used discrete wavelet transform (DWT), wavelet frame transform (WFT), and wavelet packet transform (WPT) for feature extraction. They analyze the efficiency of their approach using the count of pixels in each cluster. Khalid et al. [25] use the intensity normalization, contrast limited adaptive histogram equalization (CLAHE) as an image preprocessing technique and for classification, they used three-layer CNN. They achieved 96.64% testing accuracy with the MobileNet_v2 model. Dimpy et al. [26] used DensNets-based feature extraction technique and supervised classifier algorithm. They achieved the highest AUC score which is 0.8002.

Recently many researchers came out to detect COVID-19 infection, Dilbag et al. [27] developed a multi-objective differential evolution-based CNN algorithm to classify CT images of COVID-19. Their proposed method outperforms existing techniques with 1.9789% of accuracy, 2.0928% of F-measure, 1.8262% of sensitivity, 1.6827% of specificity, and 1.9276% of kappa statistics. Halgurd et al. [28] developed deep learning techniques using CNN to diagnose COVID-19 using CT and X-ray images. They collected CT and X-ray images from different sources such as GitHub, Redopedia, Kaggle, and BSTI. They used a transfer learning-based approach to train CNN with 16 filters. They achieved 98% accuracy using a pre-trained network and 94.1% accuracy by using their modified network. To classify COVID-19, pneumonia, and normal lungs from CT scan images, Ali et al. [29] developed long short-term memory (LSTM)-based classifier, and they used entropy feature extraction. They got the highest accuracy of 99.68% with Q-deformed entropy with LSTM. Chowdhury et al. [30] perform different experiments on existing models such as SqueezeNet, MobileNetv2, ResNet18, Inceptionv3, ResNet101, ChexNet, DensNet201, and VGG16 with augmented data and non-augmented data. They achieved 99.70% accuracy using DensNet201 with image augmentation techniques. So, they proved that image augmentation plays a major role in the preprocessing phase. Yifan et al. [31] proposed the new domain adaption technique called COVID-DA to classify pneumonia and COVID-19 from X-ray images. This method uses

very limited samples in the target domain. They achieved a 92.98% F1 score, 88.33% recall, 98.15% precision, and 0.985 AUC. Mohammad et al. [32] concatenated the Xception network and ResNet50V2 for parallel deep feature extraction tasks and apply convolutional layers to classify the images in three classes such as pneumonia, COVID-19, and normal. They achieved 99.50% accuracy to detect COVID-19 cases. Linda et al. [33] proposed the novel technique called COVID-Net to classify the COVID-19 cases from the chest X-ray images. They used a projection–expansion–projection design pattern to develop COVID-Net architecture. They achieved 93.3% accuracy, 91% sensitivity, and 98.9% specificity for COVID-19 prediction using their proposed architecture. Li [34] proposed a deep learning technique called COVNet based on ResNet50 architecture. They used a shared weight mechanism to develop their CNN model. They achieved 90% sensitivity, 96% specificity, and 0.96 AUC for COVID-19 detection.

All existing deep learning methods to detect COVID-19 are giving good results, but it is not reaching the best extent. COVID-19 is spreading quickly via air and physical touches, so only a 1% error in the screening of COVID-19 can ignore the infected person and that person can infect thousands of other people. To stop this infection chain, a higher accurate approach with very high sensitivity and specificity is required.

3 Methodology

In this section, we introduce the new preprocessing pipeline containing image enhancement and augmentation techniques. For classification, deep convolutional neural networks perform very well with the stable dataset, but sometimes due to an unstable dataset, the deep convolutional neural network cannot learn efficiently and failed to express uncertainty, and the model gets overfit. To solve this issue, Shridhar et al. [35] proposed the Bayesian approach in the convolutional neural network since the Bayesian network is very useful to capture more uncertainty in data. We adapted the Bayesian convolutional neural network (BCNN) technique given in [35] and proposed the modified version of the AlexNet architecture with the orthogonal normalization technique (ONT) in Bayesian layers. ONT is very helpful to capture more uncertainty and faster convergence. The RMSprop optimizer was used to train the model, and we found that the RMSprop optimizer is working better than Adam with the proposed model. We show how the Bayesian layer with ONT improves the performance of the model. The full analysis of the results is mentioned in Sect. 5. In Sect. 3.1, we explained the dataset preprocessing pipeline, in Sect. 3.2, we show the design of the modified AlexNet architecture and, and in Sect. 3.3, we describe the whole architecture of the Bayesian orthogonal normalized layer.

3.1 Dataset Preprocessing

Data preprocessing is an essential part of any deep learning algorithms. The proposed preprocessing pipeline is shown in Fig. 1. The original images in all used datasets [36–39] are of variable size, three channels, and low contrast images. Histogram equalization [40] was applied to improves contrast in the image, which is helpful to find better GGO patterns to detect COVID-19. First, we convert the image from RGB (Red, Green, Blue) to HSV (Hue, Saturation, Value/Brightness) and perform histogram equalization only on the V channel. After that, we combine the H and S channels of the original image with the V channel of the processed image. The luminance encoding technique called the gamma transform [41] with constant $C = 1$ and gamma $\gamma = 2$ is also applied to enhance the dynamic range of the image.

$$y = 10^{2.4065 - 2 \log \left[\frac{\text{Round} \left(\frac{cdf(V_i) - cdf_{\min}}{p - cdf_{\min}} \right)}{255} (L - 1) \right]} \tag{1}$$

Equation (1) indicates the image enhancement pipeline of the proposed method, where p indicates the total number of pixels in the image, V_i indicates each pixel coming from the V channel of the image, L indicates the gray levels of the image, cdf indicates the cumulative distribution function, and y indicates the final output of the image enhancement techniques. Grayscale transformation is also applied to the y to bring simplicity to the image. To get more randomization in the data, we

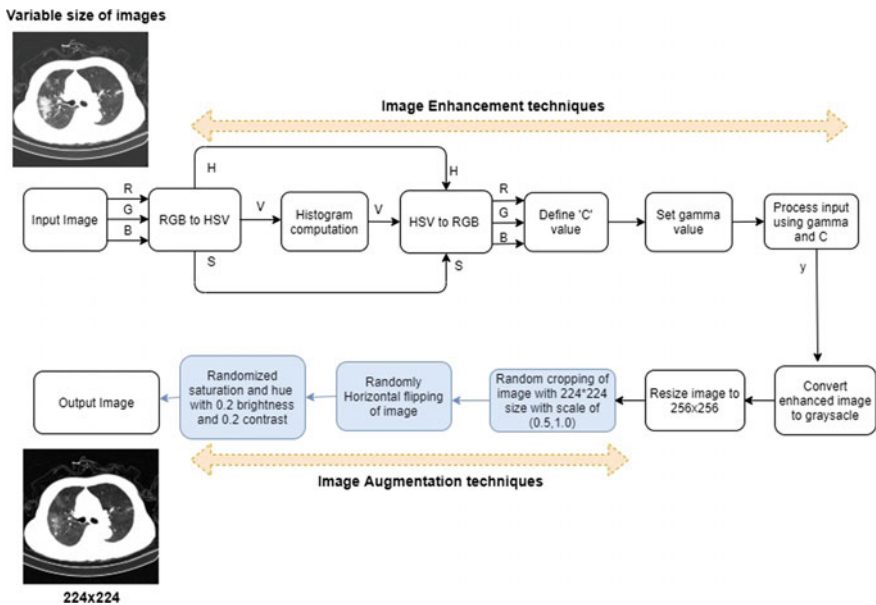


Fig. 1 Dataset preprocessing pipeline

proposed another continuous pipeline called image augmentation techniques. The image is cropped into 224×224 and applied random flipping on all input images. Furthermore, the randomized saturation and hue with 0.2 brightness and 0.2 contrast were also applied to bring more randomization in the inputs. After completion of both image enhancement and augmentation techniques, the generated image can be used for the training and testing of the model.

3.2 Modified AlexNet Architecture with Bayesian Orthogonal Normalized Layers

Figure 2 indicates the proposed modified AlexNet architecture with the Bayesian orthogonal normalized layer. The preprocessed image with size 224×224 is given to this architecture to classify it in COVID-19, pneumonia, or normal. To introduce the Bayesian inference approach in the convolutional neural network (CNN), the convolutional layer and the fully connected layer are replaced with the Bayesian orthogonal normalized layers. The whole layer architecture of the Bayesian layers with ONT is given in Sect. 3.3. We design the number of neurons in each layer and used the same dropout rate of 0.5 in the model to prevent overfitting. The softmax activation is used at the end of the model to classify the image into three classes.

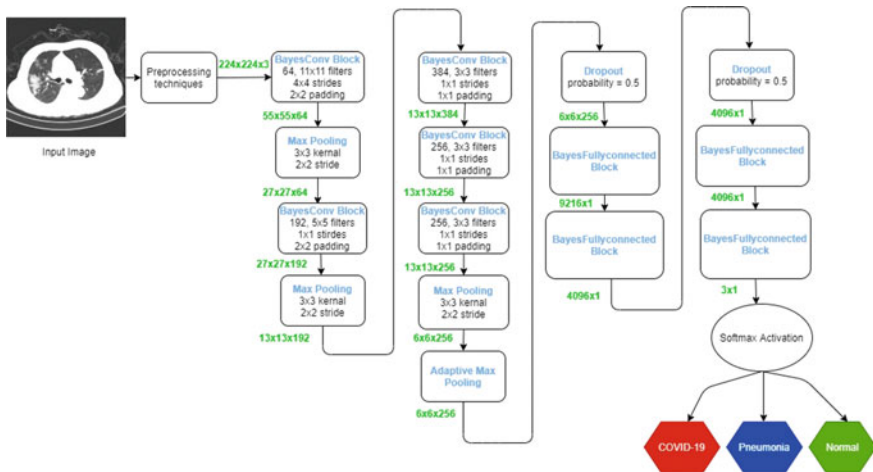


Fig. 2 Modified AlexNet architecture with Bayesian orthogonal normalized layers

3.3 Bayesian Orthogonal Normalized Layer

Simple convolutional layers can learn only a weight W and bias B . So, when the data is limited in a certain region, then the model is got overfitted. Bayesian networks [42] is a probabilistic graphical model, which is taking probabilistic decisions using conditional dependencies and capture more uncertainty. A simple Bayesian approach in CNN is proposed in [35], so we adopted their technique to create basic Bayesian layers on our proposed modified AlexNet architecture. The simple Bayesian layer in the model is not giving a stable output in all scenarios. So, we replaced all convolutional layers with the Bayesian orthogonal normalized convolutional layer, and all fully connected layers are replaced with Bayesian orthogonal normalized fully connected layer.

$$W_\sigma = \text{Log}_e(e^{W_\rho} + 1) \tag{2}$$

$$B_\sigma = \text{Log}_e(e^{B_\rho} + 1) \tag{3}$$

In Fig. 3, proposed Bayesian orthogonal normalized layer is given. The single weight and bias are replaced with some random mean μ and standard deviations σ using the normal distribution. So now, the weight is converted to the two weights called weight mean W_μ and weight standard deviations W_ρ , and the bias is converted to the two bias called the bias mean B_μ and bias standard deviations B_ρ . So, the learning parameters of the model get doubled. The W_ρ is converted to W_σ , and B_ρ is converted to B_σ using Eqs. (2) and (3), respectively, where W_σ called normalized

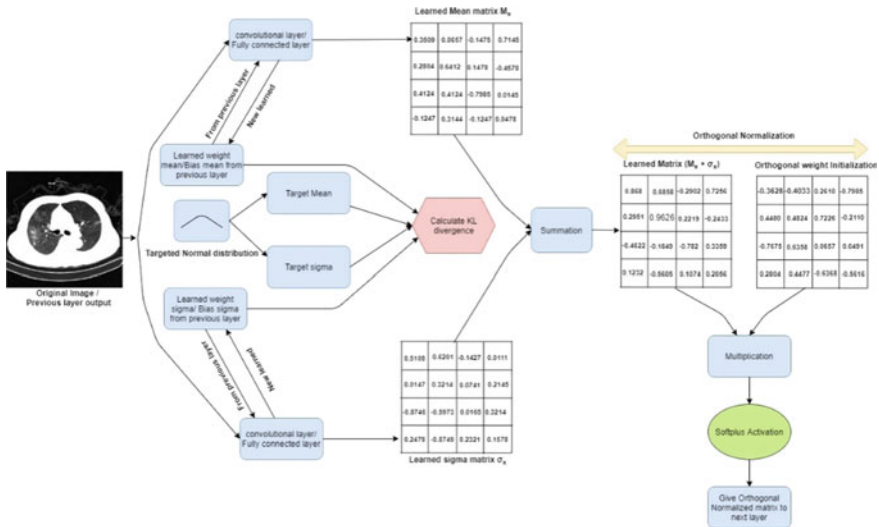


Fig. 3 Bayesian orthogonal normalized layer

weight standard deviations and B_σ called normalized bias standard deviations. So now $W_\mu, W_\sigma, B_\mu,$ and B_σ parameters are available to learn in the model. Also, for Bayesian approximation, the target mean $Target_\mu$ is set to 0, and the target standard deviation $Target_\sigma$ is set to 0.1 to approximate learned distribution.

Since the same operations are performed for both weight and bias, here we used common notations mean μ and standard deviation σ . To reduce the complexity of the model, the dropout layers with a 0.5 uniform probability is used in the model. If the current layer type is convolutional, then we used two convolutional layers to learn μ and σ for both weight and bias. The first convolutional layer will take the image I_x and the learned mean $Learned_\mu$ from the previous layer as input and give the newly learned mean $NLearned_\mu$. The second convolutional layer will take the image I_x and the learned standard deviation $Learned_\sigma$ from the previous layer as input and give the newly learned standard deviation $NLearned_\sigma$ for each feature of the input image. If the current layer type is fully connected, then we used two linear layers instead of convolutional layers to learn μ and σ for both weight and bias. To evaluate this $NLearned_\mu$ and $NLearned_\sigma$ for each feature, the Kullback–Leibler divergence (KL divergence) [43] was used as a loss function.

$$\begin{aligned}
 \text{KL} = 0.5 \sum & \left\{ 2\log\left(\frac{Target_\sigma}{NLearned_\sigma}\right) - 1 + \left(\frac{NLearned_\sigma}{Target_\sigma}\right)^2 \right. \\
 & \left. + \left(\frac{Target_\mu - NLearned_\mu}{Target_\sigma}\right)^2 \right\} \tag{4}
 \end{aligned}$$

KL divergence for this model is calculated using Eq. (4). After learning for a single layer, learned mean matrix M_x and a learned sigma matrix σ_x are generated as output for a single input image. The summation of M_x and σ_x to create the final normal distribution can be directly used for the input to the next layer. But for faster convergence, more optimization, and to improve the performance of the model, we introduce the ONT on the learned weight matrix. The ONT is based on the orthogonal initialization of the weights [44]. The QR decomposition technique [45] is used to initialize the orthogonal matrix.

$$Q, R = QR_{decomp} (N(0, 1)) \tag{5}$$

$$N_I = Softplus([M_x + \sigma_x]*[Q*S[diag(R)]]) \tag{6}$$

Equation (5) is used to find Q and R matrix using QR decomposition technique. We performed QR decomposition on a normally distributed matrix with the 0 mean and 1 standard deviation denoted by $N(0, 1)$. Equation (6) is used to find the final output of the ONT approach. In Eq. (6), M_x and σ_x denote the “learned mean matrix” and “learned sigma matrix” coming from the learned weights, $diag$ indicates the diagonal values of R matrix, S creates the matrix with the sign of $diag(R)$, and N_I indicates the input for the next layer. To initialize the matrix with semi-orthogonal

Table 1 Pseudocode for Bayesian layer with ONT

$Target_\mu = 0$ and $Target_\sigma = 0.1$
<i>Foreachlayer</i> :
$W \rightarrow W_\mu, W_\rho$ and $B \rightarrow B_\mu, B_\rho$
$W_\sigma = \text{Log}_e(e^{W_\rho} + 1)$ and $B_\sigma = \text{Log}_e(e^{B_\rho} + 1)$
(<i>Notation</i> : $W_\mu, B_\mu \rightarrow \mu$ and $W_\sigma, B_\sigma \rightarrow \sigma$)
<i>If layer == convolutional</i> :
$M_x, \sigma_x \leftarrow \begin{cases} N\text{Learned}_\mu = \text{conv1}(I_x, \text{Learned}_\mu) \\ N\text{Learned}_\sigma = \text{conv2}(I_x, \text{Learned}_\sigma) \\ KL(\text{Target}_\mu, \text{Target}_\sigma, N\text{Learned}_\mu, N\text{Learned}_\sigma) \end{cases}$
$N_l = \text{softplus}(\text{Orthogonal Normalization}(M_x + \sigma_x))$
<i>If layer == Fully connected</i> :
$M_x, \sigma_x \leftarrow \begin{cases} N\text{Learned}_\mu = \text{linear1}(I_x, \text{Learned}_\mu) \\ N\text{Learned}_\sigma = \text{linear2}(I_x, \text{Learned}_\sigma) \\ KL(\text{Target}_\mu, \text{Target}_\sigma, N\text{Learned}_\mu, N\text{Learned}_\sigma) \end{cases}$
$N_l = \text{softplus}(\text{Orthogonal Normalization}(M_x + \sigma_x))$

property [46], we make matrix Q as a uniform matrix using diagonal signs of the R matrix. This orthogonal matrix is used for normalizing the weight matrix $[M_x + \sigma_x]$. After orthogonal weight normalization, the softplus activation function is applied to ensure that σ will not become zero. If σ will become zero, then the proposed approach is the same as single weight learning. This orthogonal normalized image can be used for the input to the next layer. The whole pseudocode for the proposed Bayesian layer is given in Table 1.

4 Training Details

To perform experiments on the model, we used these [36–39] all datasets. In Sect. 4.1, a brief description of all datasets is given. In Sect. 4.2, we describe all training information such as dataset split-ups for training, loss function, optimizer, different activation function for layers, the learning rate of the model, batch size, and epochs information along with architectural detail of the training platform.

4.1 Datasets

In the initial phase of this work, we used the small CT dataset (Xingyi et al. 2020) [36] for the experiment. This dataset contains 349 COVID images and 397 non-COVID images, so it is only used for two-class classification. Since COVID-19 is a variant of pneumonia, the identification of pneumonia and COVID-19 both as a different disease would be very useful to perform the diagnosis and treatment of a large number of COVID-19 patients during the pandemic. As per our three-class classification requirements, we used the publicly available data provided by the China National Center for Bioinformation (CNCB) (Kang et al. 2020) [38]. This dataset size is nearly 14 GB, and it contains 21,395 COVID images, 36,856 pneumonia images, and 45,758 normal images.

To verify model performance on different datasets, we merge datasets from multiple sources contains X-ray and CT images. Joseph et al. [37] provide a combination of X-ray and CT images contains 856 images of COVID-19 positive patients. To add more data, [36] is also merged. To add pneumonia images, the chest X-ray dataset from Kaggle [39] is also merged, which contains 7495 pneumonia images and 1583 non-pneumonia (Normal) images. To stabilize the number of images in each class, 1589 random pneumonia images were taken from the Kaggle dataset. To add CT images for the pneumonia class, 1272 random pneumonia images were taken from CNCB [38] dataset. The final combined dataset contains 1205 COVID-19, 1980 normal, and 2861 pneumonia images.

4.2 Training Phase

Due to computation limitations, first, the model is trained on 18,000 random images from the CNCB [39] dataset. To verify the strength of the model, the data is randomly split up into different training, testing, and validation set combinations. All split-up combinations are given in Table 2. To test the effectiveness of the ONT approach, we trained the model on “without ONT” and “with ONT” approaches [47, 48]. The KL divergence is used as a loss function and RMSprop as an optimizer to train the

Table 2 Data split-ups

Cases	Training %	Training images count	Validation %	Validation images count	Testing %	Testing images count
1	50	9000	5	900	45	8100
2	50	9000	10	1800	40	7200
3	70	12,600	10	1800	20	3600
4	75	13,500	5	900	20	3600
5	75	13,500	10	1800	15	2700

model. Softplus activation is used in orthonormal Bayesian layers to make sure that sigma will never become zero at the time of the training. To perform three-class classification, softmax activation is also used in the last layer of the network. The learning rate is set to 0.0001, the batch size is 50, and every model is trained on 32 epochs for better result analysis. To load data quickly, four workers are used. The normalization of the image is also performed with the 0.485 mean and 0.229 standard deviations before giving it in the training. For the Bayesian approximation, $Target_{\mu}$ is 0, and $Target_{\sigma}$ is 0.1 for the training phase. We initialize the μ and σ for both weight and bias using the normal distribution, and it will learn in the training process. A similar configuration as above was also used to perform the training on a combined CT and X-ray images dataset.

All experiments are performed on Google Colaboratory Notebooks, which contains an Intel(R) Xeon(R) CPU @ 2.30 GHz and Tesla K80 GPU with 13 GB RAM.

5 Results and Discussion

As discussed in Sect. 4.2, first, the model is trained on 18,000 random images from CNCB [39] dataset for all split-ups given in Table 2. All the splits-ups are trained on the 32 number of epochs. To test the effectiveness of the ONT, the model is trained on “Without ONT” and “With ONT” approaches. The evaluation results of all split-ups for both approaches are given in Table 3.

Sensitivity, specificity, F1 score, and AUC score are used as an evaluation matrix, and to find these, the confusion matrix for three-class classifications is used. From the confusion matrix, the correctly classify and misclassified images are mentioned in Table 3, and we analyzed that using less training data “With ONT” approach is giving fewer (238 for “Without ONT” and 9 for “With ONT”—for case 1) numbers of miss classified images. So, it indicates that the “With ONT” approach is very useful when the dataset size is small. We also analyzed that “With ONT” approach is giving a significant improvement in sensitivity, specificity, F1 score, and AUC. From the analysis of Table 3, case 3 is giving the best results using the “With ONT” approach. The deviation in the best case (case 3) in testing accuracy of COVID-19 sensitivity, normal sensitivity, pneumonia sensitivity, COVID-19 specificity, normal specificity, pneumonia specificity, F1, AUC, misclassified images is 0.82%, 0%, 0.2%, 2.6%, 0.3%, 1.02%, 0.076%, 0.82%, 5.07%, 62, and 0.05%, 0%, 0%, 0%, 0.01%, 0.05%, 0.01%, 0.05%, 0.085%, 4 while comparing with an average of all “Without ONT” and “With ONT,” respectively. Confusion matrix and receiver operating characteristic (ROC) curve for the best case (case 3) are shown in Figs. 4 and 5, respectively.

Using case 3, we achieved 99.91% testing accuracy, 99.91% F1 score, and 100% AUC, 3597 correctly classified, and three misclassified images using the “With ONT” approach. Furthermore, we achieved 100% sensitivity, 100% specificity on COVID-19 prediction, 100% sensitivity, 99.97% specificity on normal prediction, and 100% specificity, 99.91% specificity on the pneumonia prediction.

Table 3 Obtained results for all split-up cases

Case	1		2		3		4		5	
Tech	W/o	W	W/o	W	W/o	W	W/o	W	W/o	W
TrAc	98.4	99.45	98.68	99.52	98.67	99.58	99.14	99.59	99	99.59
VAC	97.4	99.67	99.39	99.61	99.06	99.89	99.44	99.44	99.5	99.67
TeAc	97	99.88	99.11	99.87	99.80	99.91	99.66	99.88	99.8	99.81
Csen	100	100	100	100	100	100	100	100	100	100
Nsen	100	100	99	100	100	100	100	100	100	100
Psen	91	100	98	100	99	100	99	100	100	100
Cspe	99.9	100	98.83	99.95	100	100	99.87	100	99.8	100
Nspe	95.6	99.94	99.83	99.97	99.70	99.97	99.66	99.97	99.8	99.77
Pspe	100	99.88	100	99.87	100	99.91	99.96	99.87	99.9	99.97
F1	97	99.88	99.11	99.87	99.80	99.91	99.66	99.88	99.8	99.81
AUC	90.3	99.66	98.66	100	97.66	100	93	100	95	100
MC	238	9	64	9	7	3	12	4	5	5

(Tech—Technique, TrAc—Training accuracy, VAC—Validation accuracy, TeAc—Testing accuracy, Csen—COVID-19 sensitivity, Nsen—Normal sensitivity, Psen—Pneumonia sensitivity, Cspe—COVID-19 specificity, Nspe—Normal specificity, Pspe—Pneumonia specificity, MC—Misclassified images, W/o— “Without ONT”, and W— “With ONT”, Bold columns indicates the improved results in the “With ONT” (W) approach compared to the “Without ONT” (W/o) approach.)

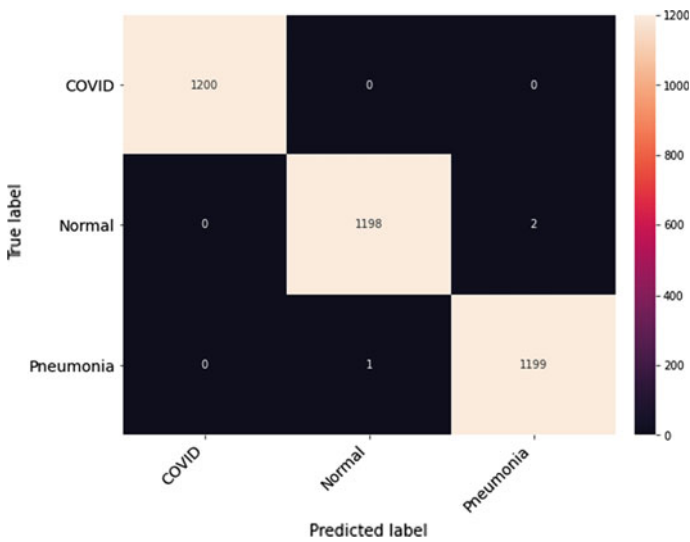
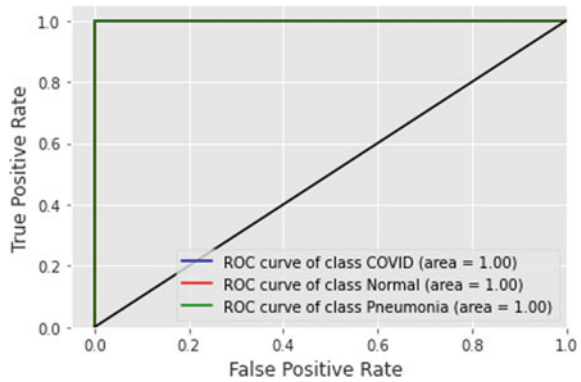


Fig. 4 Confusion matrix for case 3 (“With ONT”)

Fig. 5 ROC curve for case 3 (“With ONT”)



As discussed in Sect. 4.2, to check the efficiency of the model on other datasets, experiments with a combined (X-ray + CT) dataset is also performed, which is collected from different sources. The result analysis of this dataset is given in Table 4. From the results, we analyzed that the proposed model is giving less accuracy, sensitivity, specificity, and AUC on combined X-ray and CT images because features are in different locations for the X-ray and CT images. From 1209 testing images, 85 images are misclassified on the combined dataset. But still, it works better than existing approaches [4–7, 30]. So, we conclude that our proposed model works best with CT scan images and giving 99.91% the highest testing accuracy.

Table 4 Experimental results on combined X-ray and CT dataset on case 3 (Best case) using “With ONT”

Dataset	Combined (CT + X-ray)
TrAc	97.00
VAC	93.21
TeAc	92.96
Csen	91
Nsen	86
Psen	99
Cspe	95.96
Nspe	97.89
Pspe	94.91
F1	91.74
AUC	98.33
MC	85

(TrAc—Training accuracy, VAC—Validation accuracy, TeAc—Testing accuracy, Csen—COVID-19 sensitivity, Nsen—Normal sensitivity, Psen—Pneumonia sensitivity, Cspe—COVID-19 specificity, Nspe—Normal specificity, Pspe—Pneumonia specificity, MC—Misclassified images)

6 Conclusions

In this paper, we introduce a Bayesian convolutional neural network (BCNN) with an orthogonal normalization technique (ONT) for classification of the CT scan images in three classes of the COVID-19, normal, and pneumonia. The new pipeline is also introduced to preprocess the dataset, which consists of image enhancement techniques to improve GGO patterns and augmentation techniques. The CNCB dataset is used for the experiment, and it contains 21,395 COVID images, 36,856 pneumonia images, and 45,758 normal images. Dataset is split into train, valid, and test set in different ratios to verify the strength of the model. With less training data, the proposed approach is giving significant improvements in testing accuracy, AUC, and the number of misclassified images when compared to the other approaches. The proposed technique is giving promising results on (training%_validation%_testing%) 70_10_20 split-up with 99.91% testing accuracy with 99.91 F1 scores, 100% sensitivity, 100% specificity on COVID-19 prediction, 100% sensitivity, 99.97% specificity on normal prediction, and 100% sensitivity, 99.91% specificity on the pneumonia prediction. To verify the performance of the model on different datasets, we merged different datasets from many sources. On the combined (X-ray + CT) dataset, the proposed model is giving 92.96% testing accuracy with 91.74 F1 scores, 91% sensitivity, 95.96% specificity on COVID-19 prediction, 86% sensitivity, 97.87% specificity on normal prediction, and 99% sensitivity, 94.91% specificity on the pneumonia prediction.

The proposed model converging faster but it is taking a long time for training. So, future direction includes a reduction in training time of the proposed model using parallel processing and transfer learning approach. To diagnose COVID-19 using both X-ray and CT images, a stable combined dataset is required. Furthermore, this study can extend to find the severity of the COVID-19 disease and predict the survival time and recovery time of the patients. This will be helpful to prioritize the patients for medical treatment to save more lives. We hope that our proposed model can be helpful for a real-time quick and accurate diagnosis of the COVID-19 and pneumonia, and it will save as many lives as possible.

References

1. A.E. Gorbalenya, S.C. Baker, R.S. Baric et al., The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **5**, 536–544 (2020). <https://doi.org/10.1038/s41564-020-0695-z>
2. <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>, last accessed 2020/08/29.
3. <https://www.worldometers.info/coronavirus/>. Last accessed 2020/10/18.
4. N.W. Ohan, J.J. Heikkila, Reverse transcription-polymerase chain reaction: an overview of the technique and its applications. *Biotechnol. Adv.* **11**(1), 13–29 (1993). [https://doi.org/10.1016/0734-9750\(93\)90408-f](https://doi.org/10.1016/0734-9750(93)90408-f)

5. C. Nikam, M. Jagannath, M.M. Narayanan et al., Rapid diagnosis of Mycobacterium tuberculosis with Truenat MTB: a near-care approach. *PLoS One* **8**(1), e51121 (2013). <https://doi.org/10.1371/journal.pone.0051121>
6. T.R. Kozel, A.R. Burnham-Marusich, Point-of-care testing for infectious diseases: past, present, and future. *J. Clin. Microbiol.* **55**(8), 2313–2320 (2017). <https://doi.org/10.1128/JCM.00476-17>
7. P.C. Woo, S.K. Lau, B.H. Wong, Longitudinal profile of immunoglobulin G (IgG), IgM, and IgA antibodies against the severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein in patients with pneumonia due to the SARS coronavirus. *Clin. Diagn. Lab. Immunol.* **11**(4), 665–668 (2014). <https://doi.org/10.1128/CDLI.11.4.665-668.2004>
8. D.K. Kido, R. Gould, F. Taati, A. Duncan, J. Schnur, Comparative sensitivity of CT scans, radiographs and radionuclide bone scans in detecting metastatic calvarial lesions. *Radiology* **128**(2), 371–375 (1978). <https://doi.org/10.1148/128.2.371>
9. K. Yokota et al. (2014)., Automatic detection of GGO regions on CT images in LIDC dataset based on statistical features, in *2014 Joint 7th International Conference on Soft Computing and Intelligent Systems (SCIS) and 15th International Symposium on Advanced Intelligent Systems (ISIS)*, Kitakyushu, pp. 1374–1377. <https://doi.org/10.1109/SCIS-ISIS.2014.7044692>.
10. W. Hu, L. Xiao, J. Pennington, Provable benefit of orthogonal initialization in optimizing deep linear networks, in *8th International Conference on Learning Representations, {ICLR} 2020*, Addis Ababa, Ethiopia, April 26–30 (2020)
11. X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M. Summers, ChestX-Ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017). <https://doi.org/10.1109/cvpr.2017.369>
12. S.M. Humphries, A.M. Notary, J.P. Centeno, M.J. Strand, J.D. Crapo, E.K. Silverman, D.A. Lynch, Deep learning enables automatic classification of emphysema pattern at CT. *Radiology* **2019**(00), 1–11 (2019)
13. C. Walker, S. Gupta, R. Hartley, C.E. Brightling, Computed tomography scans in severe asthma. *Curr. Opin. Pulm. Med.* **18**(1), 42–47 (2012). <https://doi.org/10.1097/mcp.0b013e32834db255>
14. L.Y.W. Tang, H.O. Coxson, S. Lam, J. Leipsic, R.C. Tam, D.D. Sin, Towards large-scale case-finding: training and validation of residual networks for detection of chronic obstructive pulmonary disease using low-dose CT. *Lancet Digital Health* (2020). [https://doi.org/10.1016/s2589-7500\(20\)30064-9](https://doi.org/10.1016/s2589-7500(20)30064-9)
15. Z.Z. Qin, M.S. Sander, B. Rai et al., Using artificial intelligence to read chest radiographs for tuberculosis detection: a multi-site evaluation of the diagnostic accuracy of three deep learning systems. *Sci. Rep.* **9**, 15000 (2019). <https://doi.org/10.1038/s41598-019-51503-3>
16. K. Senthil Kumar, K. Venkatalakshmi, K. Karthikeyan, Lung cancer detection using image segmentation by means of various evolutionary algorithms. *Comput. Math. Methods Med.* , vol. 2019, Article ID 4909846, 16 pages (2019). <https://doi.org/10.1155/2019/4909846>
17. S. Poorna et al., Computer vision aided study for melanoma detection: a deep learning versus conventional supervised learning approach, in *Advanced Computing and Intelligent Engineering. Advances in Intelligent Systems and Computing* ed. by B. Pati, C. Panigrahi, R. Buyya, K.C. Li, vol 1082 (Springer, Singapore). https://doi.org/10.1007/978-981-15-1081-6_7
18. G.B. Saiprasath, N. Babu, J. ArunPriyan, R. Vinayakumar, V. Sowmya, K.P. Soman, Performance comparison of machine learning algorithms for malaria detection using microscopic images. *IJRAR19RP014 Int. J. Res. Anal. Rev. (IJRAR)* **6**(1) (2019)
19. M. Kiruthika, T.R. Swapna, K.C. Santhosh, K.P. Peeyush, Artery and Vein classification for hypertensive retinopathy, in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, pp. 244–248 (2019). <https://doi.org/10.1109/ICOEI.2019.8862719>
20. S.N. Kumar, D. Dinesh, T. Siddharth, S. Ramkumar, S. Nikhill, R. Lavanya, Selection of features using Particle Swarm Optimization for microaneurysm detection in fundus images, in *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, Chennai, pp. 140–144 (2017). <https://doi.org/10.1109/WiSPNET.2017.8299735>

21. M.F. Hashmi, S. Katiyar, A.G. Keskar, N.D. Bokde, Z.W. Geem, Efficient pneumonia detection in chest X-ray images using deep transfer learning. *Diagnostics (Basel)* **10**(6), 417 (2020). <https://doi.org/10.3390/diagnostics10060417>
22. S.S.K. Chouhan, V. A. Khamparia, D. Gupta, P. Tiwari, C. Moreira, R. Damaševičius, V.H.C. de Albuquerque, A novel transfer learning based approach for pneumonia detection in chest X-ray images. *Appl. Sci.* **10**, 559 (2020). <https://doi.org/10.3390/app10020559>
23. G. Verma, S. Prakash, Pneumonia classification using deep learning in healthcare. *Int. J. Innov. Technol. Exploring Eng. (IJITEE)* **9**(4) ISSN: 2278-3075 (2020). <https://doi.org/10.35940/ijitee.D1599.029420>
24. N. Parveen, R. Shabnam, M.M. Sathik, Detection of pneumonia in chest X-ray images. *J. X-Ray Sci. Technol.* **19**, 423–428, IOS Press (2011). <https://doi.org/10.3233/XST-2011-0304>
25. K. El Asnaoui, Y. Chawki, A. Idri, *Automated Methods for Detection and Classification Pneumonia based on X-Ray Images Using Deep Learning* (2020). [arXiv:2003.14363](https://arxiv.org/abs/2003.14363)
26. D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan, A. Mittal, Pneumonia detection using CNN based feature extraction, in *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Coimbatore, India, pp. 1–7 (2019). <https://doi.org/10.1109/ICECCT.2019.8869364>
27. D. Singh, V. Kumar, K.M. Vaishali, Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks. *Eur. J. Clin. Microbiol. Infect. Dis.* **39**(7), 1379–1389 (2020). <https://doi.org/10.1007/s10096-020-03901-z>
28. H.S. Maghdid, A.T. Asaad, K.Z. Ghafoor, A.S. Sadiq, M.K. Khan, *Diagnosing COVID-19 Pneumonia from X-Ray and CT Images using Deep Learning and Transfer Learning Algorithms* (2020). [arXiv:2004.00038](https://arxiv.org/abs/2004.00038)
29. A.M. Hasan, M.M. AL-Jawad, H.A. Jalab, H. Shaiba, R.W. Ibrahim, A.R. AL-Shamasneh, Classification of Covid-19 coronavirus, pneumonia and healthy lungs in CT scans using Q-deformed entropy and deep learning features. *Entropy* **22**, 517 (2020). <https://doi.org/10.3390/e22050517>
30. M.E.H. Chowdhury et al., Can AI help in screening viral and COVID-19 pneumonia? *IEEE Access* **8**, 132665–132676 (2020). <https://doi.org/10.1109/ACCESS.2020.3010287>
31. Y. Zhang, S. Niu, Z. Qiu, Y. Wei, P. Zhao, J. Yao, J. Huang, Q. Wu, M. Tan, *COVID-DA: Deep Domain Adaptation from Typical Pneumonia to COVID-19* (2020). [arXiv:2005.01577](https://arxiv.org/abs/2005.01577)
32. M. Rahimzadeh, A. Attar, A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. *Informatics Med.* **19**, 100360 (2020). <https://doi.org/10.1016/j.jimu.2020.100360>
33. L. Wang, A. Wong, *COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images* (2020). [arXiv:2003.09871](https://arxiv.org/abs/2003.09871)
34. L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, K. Cao, D. Liu, G. Wang, Q. Xu, X. Fang, S. Zhang, J. Xia, J. Xia, Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. *Radiology* **296**(2) (2020). <https://doi.org/10.1148/radiol.2020200905>
35. K. Shridhar, F. Laumann, M. Liwicki, *A Comprehensive guide to Bayesian Convolutional Neural Network with Variational Inference* (2019). [arXiv:1901.02731](https://arxiv.org/abs/1901.02731)
36. W. Gander, Algorithms for the qr-decomposition. Seminar fuer angewandte Mathematik Eidgenössische Technische Hochschule ch-8092 Zuerich (1980). Research report no. 80–02
37. A.M. Saxe, J.L. McClelland, S. Ganguli, Exact Solutions to the Nonlinear Dynamics of Learning in Deep Linear Neural Networks (2014). [arXiv:1312.6120](https://arxiv.org/abs/1312.6120)
38. X. He, X. Yang, S. Zhang, J. Zhao, Y. Zhang, E. Xing, P. Xie, *Sample-Efficient Deep Learning for COVID-19 Diagnosis Based on CT Scans*, medRxiv (2020). <https://github.com/UCSD-A14H/COVID-CT>. <https://doi.org/10.1101/2020.04.13.2006394>
39. J.P. Cohen, P. Morrison, L. Dao, K. Roth, T.Q. Duong, M. Ghassemi, *COVID-19 Image Data Collection: Prospective Predictions are the Future* (2020). [arXiv:2006.11988](https://arxiv.org/abs/2006.11988), <https://github.com/ieee8023/covid-chestxray-dataset>
40. O. Patel, Y.P.S. Maravi, S. Sharma, A comparative study of histogram equalization based image enhancement techniques for brightness preservation and contrast enhancement. *Signal Image Process. Int. J. (SIPIJ)* **4**(5) (2013). <https://doi.org/10.5121/sipij.2013.4502>

41. S. Rahman, M.M. Rahman, M. Abdullah-Al-Wadud et al., An adaptive gamma correction for image enhancement. *J. Image Video Proc.* **35** (2016). <https://doi.org/10.1186/s13640-016-0138-1>
42. D. Heckerman, A tutorial on learning with Bayesian networks, in *Innovations in Bayesian Networks. Studies in Computational Intelligence* ed. by D.E. Holmes, L.C. Jain, vol 156 (Springer, Heidelberg). https://doi.org/10.1007/978-3-540-85066-3_3
43. C. Blundell, J. Cornebise, K. Kavukcuoglu, D. Wierstra, *Weight Uncertainty in Neural Networks* (2015). [arXiv:1505.05424](https://arxiv.org/abs/1505.05424)
44. W. Hu, L. Xiao, J. Pennington, *Provable Benefit of Orthogonal Initialization in Optimizing Deep Linear Networks* (2020). [arXiv:2001.05992](https://arxiv.org/abs/2001.05992)
45. S. Mandal, S. Biswas, V.E. Balas, R.N. Shaw, A. Ghosh, Motion prediction for autonomous vehicles from lyft dataset using deep learning, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)* 30–31 Oct. 2020, pp. 768–773 (2020). <https://doi.org/10.1109/ICCCA49541.2020.9250790>
46. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering* ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh (eds.) *Lecture Notes in Electrical Engineering*, vol. 661 (Springer, Singapore) (2021). https://doi.org/10.1007/978-981-15-4692-1_58
47. K. Zhang, X. Liu, J. Shen, J. He, T. Lin, W. Li, G. Wang et al., *Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography* (2020). <https://ncov-ai.big.ac.cn/download/> <https://doi.org/10.1016/j.cell.2020.04.045>
48. <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>. Last accessed 2020/10/17

Classification of Chest Diseases from X-ray Images on the CheXpert Dataset



Hasan Nabeel Saleem, Usman Ullah Sheikh, and Saifulnizam Abd. Khalid

Abstract This work proposes a method to classify tuberculosis (TB) disease in a chest radiograph using convolutional neural network (CNN) algorithms. The main contribution of this work is to detect and classify TB disease in addition to the other five different diseases. This is achieved by using a transfer learning technique that utilizes a pre-trained CNN network to classify the TB disease. A comprehensive verification using TensorFlow is carried out to train and validate the proposed technique. This work aimed to use different pre-trained models on the CheXpert dataset and compare the area under the curve (AUC) between the CNN models. From the simulations, it was found that it is possible to classify the TB disease in addition to the other five diseases without having a degradation in the accuracy. The results confirm that transfer learning technique is superior to other methods, which exhibits less time for training and validating the datasets, and has good performance. This work achieved excellent performance in classifying three different diseases (atelectasis, edema, and tuberculosis) with AUC of 0.912, 0.945, and 0.954, respectively. Also, this work achieved second-best performance for classifying pleural effusion and consolidation diseases with AUC of 0.928 and 0.917, respectively. The method proposed in this work can be used for classification of diseases in chest radiograph as an early diagnosis tool in a clinical environment.

Keywords Deep Learning · Convolutional Neural Network · Tuberculosis · Transfer Learning · CheXpert

1 Introduction

Tuberculosis (TB) is a common disease that is caused by a specific bacteria known as bacillus *Mycobacterium tuberculosis*. In 2015, the World Health Organization has reported that around 9.6 million people were infected with TB, leading in 1.5 million

H. N. Saleem · U. U. Sheikh (✉) · S. Abd. Khalid
School of Electrical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Johor, Malaysia
e-mail: usman@utm.my

deaths [1]. The percentage of infected individuals increased dramatically by 2016: 10.4 million confirmed cases of the illness and 1.8 million deaths were recorded [2]. Most of these deaths might have been avoided if the disease was identified in earlier phases.

Chest radiographs seem to be the most popular radiological examinations. They are important for the management of different pathologies related to high death rates and present a large variety of potential knowledge, many of which is overt. The most popular studies in chest x-rays involve about the lung and heart. Thus, most research in computer-aided detection and diagnosis of chest x-rays has concentrated on these pathologies [3]. Automated thoracic radiography interpretation at the stage of performing clinicians can provide significant benefits in several clinical settings, from enhanced workflow prioritization and clinical decision-making support.

In past years, deep learning strategies have accomplished excellent results in a wide range of machine learning activities [4]. Convolutional neural networks (CNNs) have confirmed to be particularly strong for image classification tasks and have been successfully applied in galaxy morphology estimation [5], the advancement of photo-guided autonomous cars [6], face detection [7, 8], huge-scale video classification [9], and many others [10–12]. There are already several computer-aided diagnostic (CAD) systems that use CNNs to detect diseases [13–20]. However, its implementation of tuberculosis (TB) detection stays limited.

In the USA, the percentage of pathologists as the number of the medical workforce is declining [21], and the geographical distribution of pathologists favors wider, more urban counties [22]. Delays and backlogs in the timely interpretation of radiography have shown a decreased quality of healthcare in this kind of massive health institutions such as in the UK [23] and the USA [24]. The scenario is much worse in countries with low income where radiological facilities are limited for example such in some African countries [25, 26]. Precise automated radiographic analysis has the potential to increase the efficiency of the pathologist workflow and widen expert knowledge to underserved countries.

It has been recorded that there is a proportional lack of expertise in the assessment of radiology in several common locations of TB, which may minimize screening effectiveness and work-up initiatives [27, 28]. Recently, the interest in the use of computer-aided diagnosis for the detection of TB has increased [27, 29, 30].

These days, there are several extremely accurate diagnostic techniques but unfortunately, most of them are costly. The lowest cost and most common diagnostic methods, such as sputum smear microscopy, are recorded to have sensitivity problems [31]. Another common diagnostic method uses frontal chest radiographic images, but is restricted by the need for skilled staff to independently monitor every radiography that is not present in developing countries. An effective automated technique can be used as a large-scale detection tool to screen large populations efficiently [32], potentially saving many lives.

Hence, this work proposes an enhanced deep learning convolutional neural network (CNN) model with transfer learning to improve classification accuracy for TB, lung infiltrates, catheters, pneumothorax, pleural effusion, edema, and cardiomegaly diseases.

The main contributions of this work are the classification on CheXpert dataset for different pathologies using different pre-trained CNN models (DenseNet [33], GoogleNet [34], AlexNet [35], VGGNet [36], ResNet [37], and SqueezeNet [38]) by applying transfer learning technique and the classification for the newly combined dataset that contains TB disease.

2 Related Works

The first work [39] is related to CheXpert dataset, and the researchers investigate different CNN approaches to handle the uncertainty (whether the patient has a disease or not). Also, they design a labeler to automatically detect 12 different chest diseases in addition to no findings and support devices observations by using the DenseNet121 CNN architecture. The best area under the curve (AUC) was on edema = 0.941, and the worst was on cardiomegaly = 0.854 and the other observations it was at least 0.90.

The work in [40] proposed a classification approach by converting the generative adversarial network (GAN) to a semi-supervised classifier, and the researchers found that when the labeled data is limited, the GAN achieved higher accuracy compared to CNN. For example, when the labeled data was ten, the accuracy was 51.27% and 73.08%, and when the labeled data was 100, the accuracy was 57.81% and 79.58% for the CNN and GAN, respectively. But the main limitation of the GAN method is that it cannot handle classification of multi-label chest x-ray.

In the third work [41], to handle multi-label classification for 14 different chest x-ray diseases, the researchers used cascading multiple predictions using a binary relevance approach, and they found it will improve the accuracy of deep learning models such as (DenseNet161). They also solved the problem of binary relevance (BN) approach by using pairwise error (PWE) loss approach. They used the ChestX-ray14 dataset to train their model. The best AUC that they achieved was on cardiomegaly = 0.9133, and the worst was on atelectasis = 0.7618.

The next work [42] also used the ChestX-ray14 dataset to train their model. They used a long short-term memory (LSTM)-based approach to replace the need of using pre-trained models, and they found that by designing a baseline model that ignores the label dependencies, and they can significantly outperform existing pre-trained methods. Their approach has some difficulties in learning from a labelled dataset that has an unrealistic distribution of pathologies (number of normal and abnormal cases). The best AUC that they achieved was on cardiomegaly = 0.904, and the worst was on atelectasis = 0.772, which is almost similar to the previous work [41].

Attention guided convolution neural network (AG-CNN) approach was proposed in [43]. They used pre-trained CNN models (DenseNet121 and ResNet50) as a backbone, then combined it with the local cues. By doing this, they solved the problem of the poor alignment of some CXR images. One problem with their approach is that it cannot tackle the difficulties in sample collection and annotation. The best AUC that they achieved was on cardiomegaly = 0.939 which it is one of the highest AUC

to classify this disease, and the worst was on atelectasis = 0.853, this work achieved greater AUC compared with the previous works that using the same dataset with different approaches [41, 42].

Next, in [44], they used the transfer learning technique as weights initializer on the pre-trained CNN model (DenseNet121) on the ChestX-ray14 dataset. They found that one of the limitations of the ChestX-ray14 dataset is it has only a frontal view, and it has shown that up to 15% of accurate diagnoses demand lateral view. The best AUC that they achieved was on cardiomegaly = 0.9248, and the worst was on consolidation = 0.7901. Classification using a multi-label DCNN model to detect one or multiple pathologies in X-ray image was presented in [45]. The researchers found that it can be easy to detect eight common chest diseases by using unified weakly supervised multi-label image classification. They also produced a large dataset (ChestX-ray8). This dataset contains 108,948 frontal view X-ray images taken from 32,717 patients. The best AUC was on cardiomegaly with 0.8141, and the worst was on mass and nodule with 0.5609, and that is because diseases such as mass and nodule have a small object to detect. Therefore, the ratio was the lowest. The main limitation of this work is it cannot develop a fully automated deep learning system.

Lastly, in work [46], the researchers found that by using transfer learning technique which uses a CNN pre-trained model that is trained from a non-medical dataset such as ImageNet will lead to better performance and it can be achieved by ensemble of the activation layer decaf5, fully connected layer decaf6, and the GIST descriptor. The AUC for the different diseases were between 0.87 and 0.94. In [47], the researchers found that by using an ensemble of DCNN models such as GoogLeNet [34] and AlexNet [35] will lead to the best performance to detect the TB disease with AUC 0.99. The main limitation is it cannot detect other diseases.

In [48], they found that the ensemble of DCNN models will lead to a noticeable improvement compared with a single DCNN model, they achieved a very high performance of detecting TB with AUC of 0.94, in addition to three other diseases. Lastly, in [49], the detection of TB using transfer learning technique for three different pre-trained CNNs which are GoogLeNet, ResNet and VggNet, learned from a non-medical dataset (ImageNet) was done. The researchers tackled the problem of losing important information for identifying TB when resizing the chest X-ray images by using a sliding window approach, which converts the images into subregions. They achieved the best accuracy on Shenzhen and Montgomery datasets with 0.847 and 0.826, respectively.

3 Dataset and Methodology

In the following sections, the datasets, the method, and the assessment metrics to classify and detect five different diseases in addition to TB are described.

3.1 Datasets

In this work, the datasets that are used for experiments are the CheXpert, Shenzhen and Montgomery datasets. CheXpert dataset has been downloaded from the Stanford ML Group. Shenzhen and Montgomery datasets have been downloaded from the Open-i service of the National Library of Medicine that provides a different kind of chest X-ray datasets.

CheXpert dataset: CheXpert dataset consists of 224,316 chest radiographs (frontal and lateral views) of 65,240 patients with their label to show whether the patient is normal or abnormal (having different types of chest diseases) [39]. CheXpert dataset is divided into training and validation sets. The training set consists of 64,540 patients, and the validation set consists of 200 randomly sampled patients from the full dataset. Each patient may have more than one study, and each study may have two different views (frontal and lateral). The training and validation sets consist of 14 labels, 12 different diseases in addition to no finding and support devices. The 12 diseases that are in the CheXpert dataset with their distribution are shown in Table 1.

Guangdong Medical College, Shenzhen, China dataset and Montgomery County, MD, USA dataset: Shenzhen and Montgomery datasets contain only the tuberculosis (TB) disease and consist of 662 and 138 chest radiographs (frontal view) respectively, with the description including normal and abnormal [50].

Table 1 Number of studies which contain the 12 diseases in the training set

Disease	Positive (%)	Uncertain (%)	Negative (%)
No Findings	16,627 (8.86)	0 (0.00)	171,014 (91.14)
Enlarged Cardiom	9020 (4.81)	10,148 (5.41)	168,473 (89.78)
Cardiomegaly	23,002 (12.26)	6597 (3.52)	158,042 (84.23)
Lung Lesion	6856 (3.65)	1071 (0.57)	179,714 (95.78)
Lung Opacity	92,669 (49.39)	4341 (2.31)	90,631 (48.30)
Edema	48,905 (26.06)	111,571 (6.17)	127,165 (67.77)
Consolidation	12,730 (6.78)	23,976 (12.78)	150,935 (80.44)
Pneumonia	4576 (2.44)	15,658 (8.34)	167,407 (89.22)
Atelectasis	29,333 (15.63)	29,377 (15.63)	128,831 (68.71)
Pneumothorax	17,313 (9.23)	2663 (1.42)	167,665 (89.35)
Pleural Effusion	75,696 (40.34)	9419 (5.02)	102,526 (54.64)
Pleural Other	2441 (1.30)	1771 (0.94)	183,429 (97.76)
Fracture	7270 (3.87)	484 (0.26)	179,887 (95.87)
Support Devices	105,831 (56.4)	898 (0.48)	80,912 (43.12)

3.2 Pre-processing

First, all the images in Shenzhen and Montgomery datasets are resized to 320×320 without losing any important features of the X-ray images. Next, the images in the Shenzhen and Montgomery datasets are divided into training and validation sets to make it exactly like CheXpert dataset. For Shenzhen dataset, out of 662 X-ray images, 98 images that contain normal and abnormal labels for validation set are randomly chosen, the remaining images (564) are considered as a training set. For Montgomery dataset, out of 138 X-ray images, 33 images that contain normal and abnormal labels for validation set are randomly chosen, and the rest of the images (105) are considered as a training set. After combining the three datasets, the new dataset will consist of the 13 diseases (12 diseases from CheXpert and TB disease from Shenzhen and Montgomery).

In this work, from the CheXpert dataset, five kinds of chest illnesses have been taken into account: atelectasis, cardiomegaly, consolidation, edema, and pleural effusion. That is in addition to the TB disease from Shenzhen and Montgomery datasets.

3.3 Convolutional Neural Networks

In this work, we utilize various CNNs for the purpose of classification. The CheXpert dataset contains uncertainty labels; therefore, we used a binary mapping approach (U-Ones and U-Zeros) for the five different diseases to handle the uncertainty in the dataset. We explored several CNN models, e.g., DenseNet [33], GoogleNet [34], AlexNet [35], VGGNet [36], ResNet [37], and SqueezeNet [38]. All the CNN models were implemented in TensorFlow and have been fine-tuned using three different learning rates of 1.00E-03, 1.00E-02, and 5.00E-03.

3.4 Evaluation Metrics

The assessment metric that has been used in this work is the area under the ROC curves (AUC) which is similar to [39, 41–46, 48]. AUC-ROC curve is a performance measurement for classification problem. The AUC describes four quantities, true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN). The following are the evaluation metrics.

Sensitivity is computed as the true-positive rate that measures the success of identifying abnormal cases, as shown in Eq. (1).

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (1)$$

Specificity is computed as true-negative rate, which measures the success of not flagging normal cases as abnormal, as shown in Eq. (2).

$$\text{TNR} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (2)$$

False-positive rate is computed as $1 - \text{Specificity}$, as shown in Eq. (3).

$$\text{FPR} = \frac{\text{FP}}{\text{TN} + \text{FP}} \quad (3)$$

4 Results

In this work, we first trained and validated the CheXpert dataset using different CNN models, and then we trained and validated the new combined dataset. First, we used different VGG pre-trained models as follows (with and without batch normalization);

VGG11	VGG16
VGG11_BN	VGG16_BN
VGG13	VGG19
VGG13_BN	VGG19_BN

Next, we used the pre-trained AlexNet, GoogleNet and SqueezeNet for training and validation. Finally, we also experimented with DensNet and ResNet pre-trained CNN models with different number of layers, (DensNet 161-169-201, ResNet 50-101-152). All the networks were trained with a learning rate of $5.00\text{E-}03$. Figure 1 shows the average AUC for all the different pre-trained models that have been used in this work.

The newly combined dataset is then split into training set and validation set. 224,084 chest radiographs of the newly combined dataset are applied to train the whole layers of the pre-trained models. Another 366 x-ray images of the new dataset are used as the validation set. The combined dataset consists of the five different diseases from CheXpert dataset (atelectasis, cardiomegaly, consolidation, edema and pleural effusion), in addition to the new disease (TB) from Shenzhen and Montgomery datasets.

We chose the best three pre-trained CNN models that have been obtained previously in Fig. 1 (VGG19_BN, DenseNet201, and ResNet152) to be trained using the newly combined dataset. Figure 2 shows the difference between the average AUC of the said three CNN models of DenseNet201, VGG19_BN, and ResNet152 with three different learning rates.

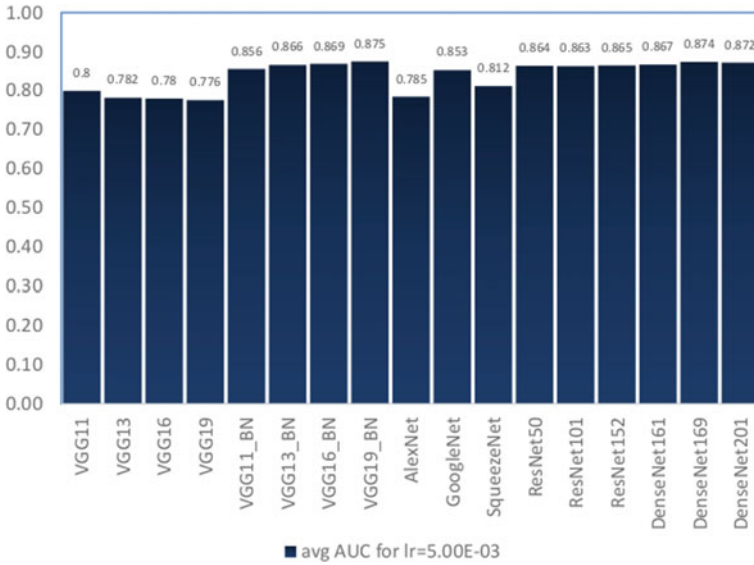


Fig. 1 Average AUC for the different pre-trained models used in this work

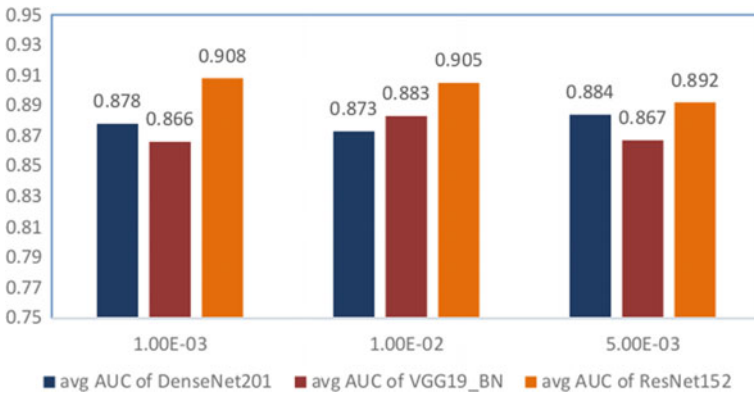


Fig. 2 Average AUCs of ResNet152, DenseNet201, and VGG19_BN pre-trained models on the newly combined dataset with three different learning rates of 1.00E-03, 1.00E-02, and 5.00E-03

5 Discussion

First, we trained and validated the different CNN models on CheXpert dataset. From Fig. 1, we can see that GoogleNet achieved the highest AUC, while the SqueezeNet CNN model achieved better performance than AlexNet model with $50 \times$ fewer parameters. Then, we did the training and the validation using DensNet and ResNet CNN models with different layers. DensNet model surpasses ResNet model with

the same learning rate. Next, we did the training and the validation on the newly combined dataset with three CNN models that achieved the best results in Fig. 1. From the average AUCs of the three pre-trained models in Fig. 2, it is easy to see that the highest AUC for the different learning rates was achieved by ResNet152. It is important to mention that the comparison of the average AUCs was between five diseases from the CheXpert dataset and six diseases (including TB disease) from the combined dataset. From Fig. 2, it is clear to conclude that the ResNet152 model outperformed the two other models.

6 Benchmarking with Existing Works

AUC of existing transfer learning CNN models in classifying different diseases is shown in Table 2. This work computes the AUC of each class across the six diseases with three different learning rates (lr) for the three different models that have been used in this work. The best results of each column are highlighted in **bold**. Table 3 shows the mean of the various methods.

This work achieved an average AUC = 0.908 for classifying six diseases (including TB) which is the highest reported and an average AUC = 0.9082 for classifying five diseases (without TB). This work applied transfer learning and exceeded the previous state-of-the-art results from the work proposed in CheXpert [39] by 0.0026 for the average AUC for five diseases (without including TB disease). AUC scores for atelectasis and cardiomegaly surpass [39] by about 0.054 and 0.03, respectively. AUC scores of edema is higher than [39] by 0.004.

When compared to [44], our work achieved better AUC scores for atelectasis, pleural effusion, consolidation and edema by about 0.1, 0.064, 0.127, and 0.066, respectively. Except for cardiomegaly, [44] reported higher AUC scores by about 0.04. Same goes for the works in [41–43], whereby our work is better for four diseases (atelectasis, pleural effusion, consolidation, and edema) with AUC scores between 0.15 and 0.021 for the different pathologies, only for cardiomegaly [41–43] achieved higher AUC scores by about 0.02, 0.03, and 0.055, respectively. When compared to [45], our work achieved better results for all five diseases (atelectasis, cardiomegaly, pleural effusion, consolidation and edema) with AUC scores of 0.196, 0.077, 0.144, 0.209, and 0.11, respectively.

For classifying TB comparing this work with [48] that had classification for four different diseases including TB (atelectasis, cardiomegaly, pleural effusion, and TB), this work got achieved higher AUC scores for (atelectasis, pleural effusion and tuberculosis) by about 0.042, 0.038, and 0.014, respectively.

Overall, this work achieved the highest performance for classifying atelectasis and edema and achieved a new state-of-the-art performance for classifying TB and at the same time achieved the second-best performance for classifying pleural effusion and consolidation.

Table 2 AUC of existing transfer learning CNN models in classifying atelectasis, cardiomegaly, pleural effusion, consolidation, edema and tuberculosis diseases

Method	CNN	Atelectasis	Cardiomegaly	Pleural Effusion	Consolidation	Edema	Tuberculosis
Irvin et al. [39]	D-121	0.858	0.854	0.936	0.939	0.941	-
Rajpurkar et al. [44]	D-121	0.809	0.924	0.863	0.790	0.888	-
Yao et al. [42]	D-/	0.772	0.904	0.859	0.788	0.882	-
Wang et al. [45]	R-50	0.716	0.807	0.784	0.708	0.835	-
Islam et al. [48]	R-50-101-152/ V-16-19/AlexNet	0.870	0.940	0.890	-	-	0.940
Kumar et al. [41]	D-161	0.761	0.913	0.863	0.783	0.888	-
Guan et al. [43]	D-121 / R-50	0.853	0.939	0.904	0.842	0.924	-
Bar et al. [46]	Decaf (GIST + L5 + L6)	-	0.940	0.920	-	-	-
This work	R-152 lr = 1E-3	0.910	0.852	0.916	0.882	0.934	0.954
	R-152 lr = 1E-2	0.892	0.884	0.928	0.892	0.945	0.890
	R-152 lr = 5E-3	0.883	0.823	0.899	0.914	0.916	0.917
	D-201 lr = 1E-3	0.874	0.800	0.903	0.856	0.935	0.902
	D-201 lr = 1E-2	0.868	0.845	0.899	0.849	0.903	0.874
	D-201 lr = 5E-3	0.878	0.855	0.918	0.834	0.941	0.878
	V-19 lr = 1E-3	0.855	0.821	0.875	0.869	0.901	0.875
	V-19BN lr = 1E-2	0.912	0.786	0.900	0.917	0.940	0.845
	V-19BN lr = 5E-3	0.892	0.807	0.882	0.895	0.931	0.798

Table 3 Average AUC of each method across the six diseases

Method	Mean/2 diseases (Cardiomegaly, Pleural Effusion)	Mean/4 diseases (Atelectasis, Cardiomegaly, Pleural Effusion and TB)	Mean/5 diseases All five diseases without TB	Mean/6 diseases All five diseases with TB	
Irvin et al. [39]	0.895	–	0.905	–	
Rajpurkar et al. [44]	0.894	–	0.855	–	
Yao et al. [42]	0.881	–	0.814	–	
Wang et al. [45]	0.795	–	0.770	–	
Islam et al. [48]	0.915	0.910	–	–	
Kumar et al. [41]	0.888	–	0.842	–	
Guan et al. [43]	0.920	–	0.892	–	
Bar et al. [46]	0.930	–	–	–	
This work	R-152 lr = 1E-3	0.884 0.906	0.908 0.898	0.898 0.908	0.908 0.905
	R-152 lr = 1E-2	0.865	0.880	0.887	0.892
	R-152 lr = 5E-3				
	D-201 lr = 1E-3	0.851 0.872	0.869 0.871	0.873 0.872	0.878 0.873
	D-201 lr = 1E-2	0.886	0.882	0.885	0.884
	D-201 lr = 5E-3				
	V-19 lr = 1E-3	0.848 0.843	0.856 0.860	0.864 0.891	0.866 0.883
	V-19BN lr = 1E-2				
	V-19BN lr = 5E-3	0.844	0.844	0.881	0.867

7 Conclusions

In this work, a newly combined dataset was created based on three chest X-ray datasets which includes all the chest diseases and also TB. We have demonstrated the performance of various pre-trained CNNs with different learning rates, batch normalization, and various network layers. The highest average AUC that has been achieved in this work is 0.875 and 0.874 using VGGNet and DensNet models, respectively. This work achieved excellent performance in classifying three different diseases

(atelectasis, edema, and tuberculosis) with AUC 0.912, 0.945, and 0.954, respectively, and second-best performance for classifying pleural effusion and consolidation with AUC 0.928 and 0.917, respectively. This work can further be improved by trying with other CNNs such as SE-ResNet and also by increasing the types of chest diseases it can classify.

Acknowledgements The authors thank the Ministry of Education Malaysia and Universiti Teknologi Malaysia (UTM) for their support under the Research University Grant (GUP), grant number Q.J130000.2523.19H45.

References

1. W. H. Organization and W. H. Organization, *Global Tuberculosis Report 2015* (2015)
2. D. Falzon et al., World Health Organization treatment guidelines for drug-resistant tuberculosis, 2016 update. *Eur. Respir. J.* **49**(3), 1602308 (2017)
3. B. van Ginneken, L. Hogeweg, M. Prokop, Computer-aided diagnosis in chest radiography: beyond nodules. *Eur. J. Radiol.* **72**(2), 226–230 (2009)
4. J. Schmidhuber, Deep learning in neural networks: an overview. *Neural Netw.* **61**, 85–117 (2015)
5. S. Dieleman, K.W. Willett, J. Dambre, Rotation-invariant convolutional neural networks for galaxy morphology prediction. *Mon. Not. R. Astron. Soc.* **450**(2), 1441–1459 (2015)
6. B. Huval et al., *An Empirical Evaluation of Deep Learning on Highway Driving*. arXiv Prepr. arXiv1504.01716 (2015)
7. H. Li, Z. Lin, X. Shen, J. Brandt, G. Hua, A convolutional neural network cascade for face detection, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5325–5334 (2015)
8. S. S. Farfadi, M.J. Saberian, L.-J. Li, Multi-view face detection using deep convolutional neural networks, in *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, pp. 643–650 (2015)
9. A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, L. Fei-Fei, Large-scale video classification with convolutional neural networks, in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 1725–1732 (2014)
10. J. Zbontar, Y. LeCun, Computing the stereo matching cost with a convolutional neural network, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1592–1599 (2015)
11. S. Ji, W. Xu, M. Yang, K. Yu, 3D convolutional neural networks for human action recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(1), 221–231 (2013)
12. S. Sudholt, G.A. Fink, PHOCNet: a deep convolutional neural network for word spotting in handwritten documents, in *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 277–282 (2016)
13. G. Apou et al., Detection of lobular structures in normal breast tissue. *Comput. Biol. Med.* **74**, 91–102 (2016)
14. K.-L. Hua, C.-H. Hsu, S.C. Hidayati, W.-H. Cheng, Y.-J. Chen, Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *Onco. Targets. Ther.* **8** (2015)
15. M. Havaei et al., Brain tumor segmentation with deep neural networks. *Med. Image Anal.* **35**, 18–31 (2017)
16. J. Kawahara, A. BenTaieb, G. Hamarneh, Deep features to classify skin lesions, in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 1397–1400 (2016)

17. B. Van Ginneken, A.A.A. Setio, C. Jacobs, F. Ciompi, Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans, in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pp. 286–289 (2015)
18. Y. Miki et al., Classification of teeth in cone-beam CT using deep convolutional neural network. *Comput. Biol. Med.* **80**, 24–29 (2017)
19. F. Ciompi et al., Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box. *Med. Image Anal.* **26**(1), 195–202 (2015)
20. Y. Bar, I. Diamant, L. Wolf, H. Greenspan, Deep learning with non-medical training used for chest pathology identification, in *Medical Imaging 2015: Computer-Aided Diagnosis*, vol. 9414, p. 94140V (2015)
21. A.B. Rosenkrantz, D.R. Hughes, R. Duszak Jr., The US radiologist workforce: an analysis of temporal and geographic variation by using large national datasets. *Radiology* **279**(1), 175–184 (2015)
22. A.B. Rosenkrantz, W. Wang, D.R. Hughes, R. Duszak Jr., A county-level analysis of the US radiologist workforce: physician supply and subspecialty characteristics. *J. Am. Coll. Radiol.* **15**(4), 601–606 (2018)
23. A. Rimmer, *Radiologist Shortage Leaves Patient Care at Risk, Warns Royal College*, vol. 359 (2017)
24. S. Bastawrous, B. Carney, Improving patient safety: avoiding unread imaging exams in the national VA enterprise electronic health record. *J. Digit. Imaging* **30**(3), 309–313 (2017)
25. D.A. Rosman et al., Imaging in the land of 1000 hills: Rwanda radiology country report. *J. Glob. Radiol.* **1**(1), 5 (2015)
26. F.S. Ali, S.G. Harrington, S.B. Kennedy, S. Hussain, Diagnostic radiology in Liberia: a country report. *J. Glob. Radiol.* **1**(2), 6 (2015)
27. J. Melendez et al., An automated tuberculosis screening strategy combining X-ray-based computer-aided detection and clinical information. *Sci. Rep.* **6**, 25265 (2016)
28. A.H. Van't Hoog et al., High sensitivity of chest radiograph reading by clinical officers in a tuberculosis prevalence survey. *Int. J. Tuberc. Lung Dis.* **15**(10), 1308–1314 (2011)
29. S. Antani, S. Candemir, Automated detection of lung diseases in chest X-rays. *US Natl. Libr. Med.* (2015)
30. S. Jaeger et al., Automatic screening for tuberculosis in chest radiographs: a survey. *Quant. Imaging Med. Surg.* **3**(2), 89 (2013)
31. C.C. Leung, Reexamining the role of radiography in tuberculosis case finding. *Int. J. Tuberc. Lung Dis. Off. J. Int. Union against Tuberc. Lung Dis.* **15**(10), 1279 (2011)
32. S. Jaeger et al., Automatic tuberculosis screening using chest radiographs. *IEEE Trans. Med. Imaging* **33**(2), 233–245 (2014)
33. S. Jégou, M. Drozdal, D. Vazquez, A. Romero, Y. Bengio, The one hundred layers tiramisu: fully convolutional densenets for semantic segmentation, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 11–19 (2017)
34. C. Szegedy et al., Going deeper with convolutions, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9 (2015)
35. A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in *Advances in Neural Information Processing Systems*, pp. 1097–1105 (2012)
36. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. *arXiv Prepr. arXiv1409.1556* (2014)
37. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778 (2016)
38. F.N. Iandola, S. Han, M.W. Moskewicz, K. Ashraf, W.J. Dally, K. Keutzer, SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv Prepr. arXiv1602.07360* (2016)
39. J. Irvin et al., *CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison*. *arXiv Prepr. arXiv1901.07031* (2019)

40. A. Madani, M. Moradi, A. Karargyris, T. Syeda-Mahmood, Semi-supervised learning with generative adversarial networks for chest X-ray classification with ability of data domain adaptation, in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 1038–1042 (2018)
41. P. Kumar, M. Grewal, M.M. Srivastava, Boosted cascaded convnets for multi-label classification of thoracic diseases in chest radiographs, in *International Conference Image Analysis and Recognition*, pp. 546–552 (2018)
42. L. Yao, E. Poblenz, D. Dagunts, B. Covington, D. Bernard, K. Lyman, *Learning to Diagnose from Scratch by Exploiting Dependencies Among Labels*. arXiv Prepr. arXiv1710.10501 (2017)
43. Q. Guan, Y. Huang, Z. Zhong, Z. Zheng, L. Zheng, Y. Yang, *Diagnose Like a Radiologist: Attention Guided Convolutional Neural Network for Thorax Disease Classification*. arXiv Prepr. arXiv1801.09927 (2018)
44. P. Rajpurkar et al., *CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning*, pp. 3–9 (2017)
45. X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M. Summers, ChestX-Ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. *IEEE Conf. Comput. Vis. Pattern Recogn. (CVPR)* **2017**, 3462–3471 (2017)
46. Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, H. Greenspan, Chest pathology detection using deep learning with non-medical training, in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pp. 294–297 (2015)
47. P. Lakhani, B. Sundaram, Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology* **284**(2), 574–582 (2017)
48. M.T. Islam, M.A. Aowal, A.T. Minhaz, K. Ashraf, *Abnormality Detection and Localization in Chest X-Rays Using Deep Convolutional Neural Networks*. arXiv Prepr. arXiv1705.09850 (2017)
49. U.K. Lopes, J.F. Valiati, Pre-trained convolutional neural networks as feature extractors for tuberculosis detection. *Comput. Biol. Med.* **89**(February), 135–143 (2017)
50. S. Jaeger, S. Candemir, S. Antani, Y.-X.J. Wang, P.-X. Lu, G. Thoma, Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quant. Imaging Med. Surg.* **4**(6), 475 (2014)

Feature-Based Overview of Online Comments of Web-Based Healthcare Products



Saroj Kushwah, Bharti Kalra, and Sanjoy Das

Abstract The increasing accessibility and prevalence of opinion-rich tools, such as the review sites for the healthcare products offered online, render it impossible for consumers to select the best product from a vast range of items. The amount of consumer opinions accessible for fashionable items can be thousands. Both reviews are challenging for consumers to interpret, and if they read just a handful of these reviews, they will have a selective perception of the product. Product manufacturers may often have trouble maintaining, monitoring, and recognizing the consumers' opinions on the goods. Several study works have in the past been suggested for solving these problems, but they have several limitations: the structures introduced are entirely invisible and the feedback are not so easy to perceive and require longer to evaluate since, aside from individual feature opinions, the function-based description system applied is broader than those used for examination. Here, we suggested a dynamic framework for the summary of consumer views on online healthcare products centered on functionality that works according to the product domain. Each period after extraction, we conduct the following work: Initially, the recognition of the characteristics of the product from the views of customers is carried out. Their respective views are then derived for each feature and their alignment or (negative/positive) polarity is identified. Finally, feature-based analytics were summarized by taking the corresponding extracts from each feature opinion and putting them in their respective cluster-based features. These type of feature-based extracts shall be absorbed easily by the end user.

Keywords Sentiment classification · Summarization · Opinion mining

S. Kushwah (✉) · B. Kalra
Noida International University, Noida, India

B. Kalra
e-mail: bharti.kalra@niu.edu.in

S. Das
Indira Gandhi National Tribal University-RCM, Imphal, India

1 Introduction

The exponential growth of e-commerce sites has greatly expanded the online purchasing of goods. Because of the broad selection of items and the ease of shopping with enticing offerings, these sites are common with consumers and even manufacturers. Around the same period, though, it is impossible for consumers to get the support of experts offering the healthcare product to purchase. One strategy used to overcome these issues is to include merchant metadata for the healthcare product offered online. And the issue with the metadata is that by only understanding the functionality of the healthcare product, the buyer considers it impossible to determine on the product, contributing to consumer insecurity and having a detrimental impact on their online e-commerce sales.

To a certain degree, the scientific group is discussing the method used to tackle this issue. Amazon, Flipkart, and Netflix, for example, use recommendation framework [1, 2]. The issue with this recommended method is that it is fully invisible, which makes it impossible for consumers to trust in the recommended goods (healthcare products). This is because there are no available information about how a customer can buy an object. The consumer should assume the algorithm below the prescribed system blindly.

1.1 Analysis Section

The analysis is done as follows: The user feedback of the commodity are originally derived from the Internet. These customer reviews shall be evaluated semantically as follows (1) Domain features of the online product are derived from customer reviews. Until removing the functions, all domain synonyms are combined into the same role category, terms, and phrases. (2) Sentences of opinion shall be defined and corresponding terms of opinion extracted. Each word of sentiments is evaluated for its positive or negative orientation [3]. As a statement may include more than one words of sentiments, current strategies have neglected to answer them properly. We created a novel method of solving this problem by taking into account any word's distance from the product function and measuring the phrase's overall opinion. This appears to be incredibly valuable. Finally, summarize the feedback by collecting extracts for each pair of feature opinions and putting them in their respective feature-based clusters. The consumer can quickly consume these feature-based extracts. Here the method deployed would be interactive, i.e., all user ratings introduced during that time are retrieved from the Internet for each period (daily or hourly), and a modified feature-based overview is created [4].

Table 1 Top seven encouraging features-based cluster for healthcare contraceptive product

Infection	250+
Effectiveness	143+
Size	130+
Cost	90+
Weight	85+
Time duration	120+
Maintenance	139+

Table 2 Top five depressing features which based cluster for healthcare contraceptive product

Cost	130–
Size	30–
Maintenance	82–
Infection	27–
Effectiveness	67–

Illustration

An illustration shows below the domain-specific feature-based overview method. Take healthcare contraceptive product example. The summary produced for contraceptive is as follows:

As seen in Table 1, the contraceptive product and scale features more favorable feedback, respectively, 250, 143, 130, 90, 85, 120, and 139 positive views of infection, effectiveness, size, cost, weight, time duration, and maintenance, respectively, and every other feature can be seen by tapping on their own features. Likewise, cost and maintenance have the most negative feedback with 130 and 82 as seen in Table 2.

However, collecting excerpts of positive and poor points from feedback is not a trivial job. We suggest in this research a method that automatically collects function-based opinions and summarizes positive and negative feature components.

We suggested a range of approaches for generating feature-based opinion overview reports for the product utilizing the knowledge available in the natural language analysis, artificial intelligence, and the opinion mining [5]. We would use the user feedback from the Amazon, Flipkart, Snapdeal healthcare product platform to test our suggested approach. Our findings indicate that these approaches are extremely successful.

2 Literature Review

Our study is closely aligned with [6, 7] in which the scholars propose that a sentence-based analysis takes the most typical features from relation rules [8]. The idea

behind this technique is that in a market study of a commodity (co-occurrence-based approach), the product characteristics more frequently exist than other meanings or words phrases. And association rules mining has many significant downsides, and association rule mining provides [9] different features that do not mirror the commodity property, but rather a variety of phrases sometimes occurring like, say, “comments” or “problems” [10]. We have therefore formed a paradigm which represents the synthesis of association law and probabilistic process. This is primarily because each product region (e.g., the contraceptive area of the features like “Infection” or “Effectiveness”) has its own language, i.e., in a document that belongs to this field, signs or nouns describing features (features are typically nouns[11]) are more likely to appear than in documents belonging to some other commodity field. Therefore, during the development process we will remove all standard descriptive phrases using the above probabilistic model.

In [12], the student used a tool for obtaining information that is the central mechanism for opinion. Device form for query answers was built with a summary opinion representation. The author recommends “picture templates” to be used as a summary illustration for the views expressed in the user analysis. This is a different mission. We cannot use a generation style definition. We aim to extract the domain features and their associated opinions so as to automatically generate a definition.

In this paper [13], authors concentrate on feature focused mining opinion and this paper mainly studies product categorization based on user numbers—review created available on the various Web sites. The first is the suggested multifunction segmentation process, which incorporates the multifunction analysis sentences in the single device. Sections of the speech and meaning details shall be utilized to judge the recognition of irrelevant feature, feeling terms are used to define the polarity of the item and eventually a K-medoid clustering shall be used for the categorization of product features.

In this study [14], authors address basic problem review classification, detection, and emotion analysis as a rather difficult challenge in several various applications and areas, while there is also been studies in this field in the past decade for review analysis, more remains to be achieved to build a method that can be successful and reliable for true existence. A shortage of tools and corpus needs to be dealt with multiple languages used in other fields than Chinese or English.

In this article [15], the author suggested an online consumer review feature-based overview in order to generate a relevant domain-specific product summary. They use the stopping and stemming approach and the fuzzy computer methodology to optimize accurateness and improve storage space as well as a high quality of knowledge with an improved precision for the retrieval of features and polarity detection.

This paper [16] addresses numerous problems and difficulties that data scientists have to address as lack of knowledge, access to the consistency and quantities of data, cleanups of dirty data processes, data protection and security issues... separated into work descriptions. Other data scientist can build adaptations, clusters and charts, implement them unregulated on multiple datasets and data styles as well as create the meta-algorithms that support data from different related datasets.

In this paper [17], authors suggested HAC methodology to categorize the field of contraceptive medicine to create a more comprehensive and appropriate overview. They have used the FP growth algorithm to classify the frequent function. The planned paper utilizes an online text dictionary, thesaurus, and WordNet to classify the opinion analyses and uses the hierarchical clustering approach to enhance the efficiency of the scheme.

3 Methodologies

In this report, we will update the work [18], which implemented a general definition opinion dependent on features. We have however modified our approach dramatically to make the framework function in line with the healthcare product area. This is where the deployed method is interactive, i.e., after any cycle (hourly or daily), all updated user input is taken from the Web.

3.1 *Data Preprocessing and Extraction of Feature*

Before the frequent features are discovered, all consumer reports are preprocessed in the following manner. (1) Fuzzy parallels the removal/replacement of miswords in the text. (2) Alliance of domain synonyms with the usage of the semi-supervised learning dilemma in [19], like, for the use of noun phrases or nouns are the domain synonyms (Fig. 1).

The feature extraction contains of two different phases: a) Speech component labeling all documentation that reflect a product's consumer feedback. b) Extraction of unique domain attributes. The term in the records is labeled with the respective speaking portion of POS marking. We used the Stanford NLP Parser [20] for the POS tagging of records, which tags all documentation and produces XML tagged POS doc. As a consequence paper.

The next step is to define unique domain characteristics. In this method, not only common features but also healthcare product-specific features are extracted. Previous opinion polling suggests that a product characteristics are typically expressed by nouns [21]. Almost every attribute is described with nouns, although not all nouns may be features with regard to the area of the device, for example, my feedback on healthcare product reviews are often correct. Comments reflect nouns here, but for the contraceptive product analysis they are not.

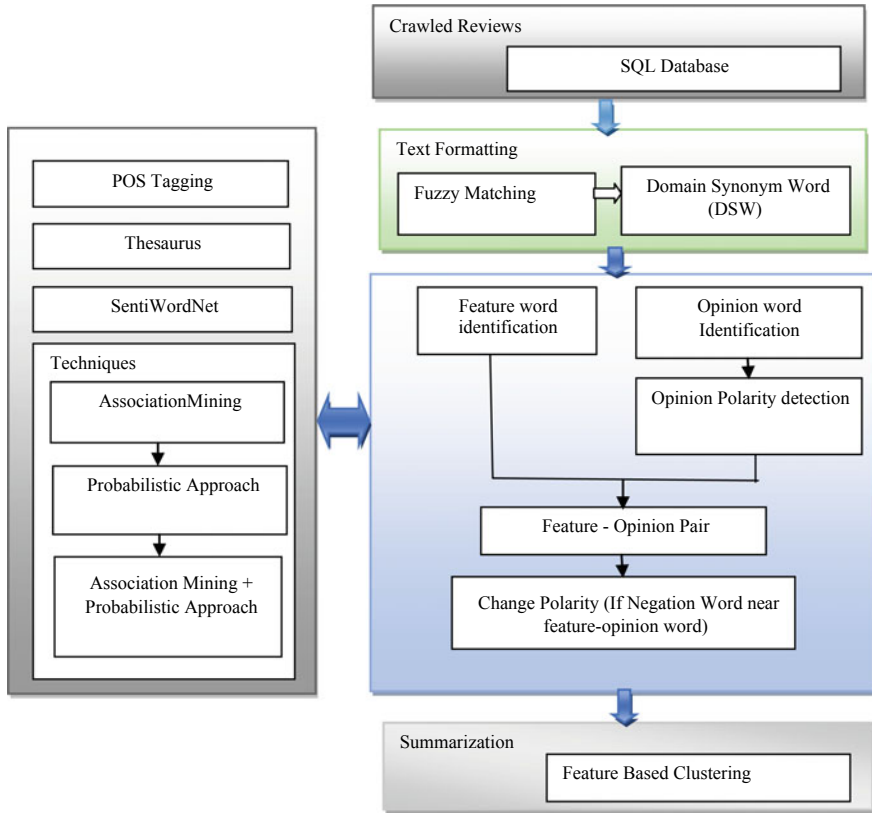


Fig. 1 A flowchart of the feature-based healthcare product summarization (FBHPS)

3.2 Extraction of Opinion Word

Extraction of opinions is done in a few different steps. (a) Extraction of all the terms of opinion and identification of polarity (positive/negative) of a single word of opinion. (b) The creation of feature opinion pairs by assigning the nearest feature to the word of opinion. (c) Ultimately, it shall be performed to evaluate the word negation near the opinion word and to detect the final polarity for all the pairs of features.

The closest attribute of the sentence is allocated to the viewpoint described in the above process. The simple reason for granting opinion to the nearest feature is that the feature’s opinion term is often the closest around it.

Although a sentence may involve more than one opinion term, current strategies have struggled to cope well with them. We created a novel method of solving this problem by taking into account any word’s distance from the product function and measuring the phrase’s overall opinion. This is quite helpful relative to the one in [22].

First, we strive to locate negation in any adjective's neighborhood. If a negative term is detected, next the pair of feature opinions is inverted and a final polarity is produced of each pair of feature opinions. Here are both negative polarities for each attribute in consumer feedback of a single product, once each feature's polarity profile is determined.

4 Classification (Feature-Based)

During this time, $2n$ clusters are generated to contain n functions extracted by the above-mentioned extraction process, each with two positive as (1) more favorable or less favorable and one negative (unfavorable) clusters by the use of classification approaches such as (1) association mining, (2) probabilistic, and (3) combination of probabilistic and association mining approaches. The helping cluster stores all healthcare products feature reviews and all unfavorable feature reviews in the bad cluster by comparing the three different classification approaches. The following is the extract of related user input extracts completed and positioned in the same application-based community of each feature opinion pair. Excerpts are applied to the clusters based on the healthcare product's function in the excerpt and on the polarity of the pair of opinions (as found in the opinion extract step above), i.e., if the excerpt is either positive or negative. Opinion extracts from each of the feature-based cluster provide a very simple and efficient overview of the features concerned.

5 Experiments

In Java, a system, known as dynamic FBS, was implemented which works in accordance with the product domain. Our consumer response tests have been carried out on three healthcare products: contraceptive products, energy products, and fitness products. Feedback from product customers was obtained from the Amazon Web site [23], Flipkart, and Snapdeal. The IDE used is the Maven Eclipse Plug-in.

The method is as below: the identical characteristics for each and every healthcare product are collected. Subsequently, their corresponding user ratings are picked for each feature with the term opinion. Texts are then derived from each of the selected function reviews (a term that basically includes the review pair). The extracts are then put on a positive or negative basis in their respective clusters. The top five extracts of the positive or negative functional cluster are checked manually in order to find that the functional analysis approach is right. The results indicate that 94% of all excerpts are accurately labeled on the contraceptive, as the other 8% are inaccurate. However, using the method of [24] quotes for correctly categorized just 89 percent. Second, the move is due to a revised process by which different competing points of view are allocated to each aspect of the product reviews. Secondly, domain feature extraction

Table 3 Feature extraction accuracy of the system

Products	Feature Extraction Accuracy (Association Mining) (%)	Feature Extraction Accuracy (Probabilistic) (%)	Feature Extraction Accuracy (Association Mining and Probabilistic Approach) (%)
Contraceptive Product	84	89	94
Energy product	86	91	96
Fitness product	78	85	90

Table 4 Opinion sentence polarity—detection accuracy of the system

Products	Opinion sentence polarity detection accuracy [13] (%)	Opinion sentence polarity detection accuracy (%)
Contraceptive product	80	82
Energy product	83	87
Fitness product	75	76

in relation to is delegated. The extracted attribute was manually checked for its exactness and the extracted feature was correct 96%, compared with 91% [25] and 86%. Accuracy development is attributed to (1) the probabilistic method used in word processing. (2) In order to extract a special domain function, we used a mixture of relationship mining and probabilistic method [26]. The initial processing [3] of type nouns or noun phrases that are the domain synonyms that must be classified within one noun community utilizing the semi-controlled learning problem was conducted before the extraction of features [27, 28]. For contraceptive, energy and fitness products, Tables 3 and 4 display the results for correct removal and polarity detection. The here embraced methodology is digital, i.e., any new customers introduced over the period are obtained from the Amazon Web site after some length of time (daily or hourly) and an updated functional overview is produced. It shows regular customer input improvements in the consistency of our operational extraction method. It also indicates changes to the quality of our opinion word polarity recognition through daily consumer feedback [29]. Our FBHPS architecture (shown in Fig. 2) can be considered to be fairly robust for regular user input updates.

The online survey performed by [30] assessed consumer interaction with the FBHPS framework. The ultimate objective was to evaluate the efficacy of analyzing the whole customer feedback with the overview created by the feature-based resuming framework. Our approach to summary generation based on features is highly powerful. In general, the precision of our feature extraction polarity detection and sentence opinion method is satisfactory.

6 Conclusion

Here, we suggested a novel methodology for the summarization of consumer feedback focused on complex characteristics that function according to the healthcare product domain. It was focused on the production of natural languages and opinion mining. The findings suggest that the strategies suggested are extremely useful and successful in carrying out their assignments. It has now become much simpler for consumers to absorb details in various healthcare product reviews by enabling users to skim in fast and productive ways across the product reviews. We would strive in future to boost the accuracy of our opinion identification polarity and the extraction algorithms. In addition to function to explore, communicate thoughts with verbs, adverbs, and substantives. The evaluations conducted in this research article present a strong foundation to build a feature-based summary framework much more effective.

References

1. G. Linden, B. Smith, J. York, Amazon.com recommendations: item-to-item collaborative filtering, in *IEEE Internet Computing*, vol. 7, no. 1, pp. 76–80, Jan.-Feb. 2003. <https://doi.org/10.1109/MIC.2003.1167344>
2. R.M. Bell, Y. Koren, C. Volinsky, All together now: a perspective on the Netflix Prize. *Chance* **23**, 24–24 (2010)
3. J. Goldstein, M. Kantrowitz, V. Mittal, J. Carbonell, Summarizing text documents: sentence selection and evaluation metric, in *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, no. 8, pp. 121–128. <https://doi.org/10.1145/312624.312665>, <https://doi.org/10.1145/312624.312665>, Berkeley, California, USA (1999)
4. G. Salton, A. Singhal, C. Buckley, M. Mitra, Automatic text decomposition using text segments and text themes, in *Proceedings of the Seventh ACM Conference on Hypertext*, pp. 53–65, no. 13. <https://doi.org/10.1145/234828.234834>. <https://doi.org/10.1145/234828.234834>, Bethesda, Maryland, USA (1996)
5. J. Tait, Automatic Summarizing of English Texts. Ph.D. Dissertation, University of Cambridge (1983)
6. M. Hu, B. Liu, Mining and summarizing customer reviews, in *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, no. 10, pp. 168–177. <https://doi.org/10.1145/1014052.1014073>. <https://doi.org/10.1145/1014052.1014073>, Seattle, WA, USA (2004)
7. B. Pang, L. Lee, Opinion mining and sentiment analysis, in *Foundations and Trends in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, Jan. 2008. <https://doi.org/10.1561/1500000011>. <https://doi.org/10.1561/1500000011>
8. R. Agrawal, R. Srikant, Fast algorithms for mining association rules in large databases, in *Proceedings of the 20th International Conference on Very Large Data Bases*, Morgan Kaufmann Publishers Inc., no. 13, pp. 487–499, San Francisco, CA, USA (1994)
9. S. Morinaga, K. Yamanishi, K. Tateishi, T. Fukushima, Mining product reputations on the Web, in *8th ACM Int. Conf. on Knowledge Discovery and Data Mining (KDD)*, no. 9, pp. 341–349, Edmonton, Canada (2002). <https://doi.org/10.1145/775047.775098>, Doi:10.1145/775047.775098
10. C. Scaffidi, K. Bierhoff, E. Chang, M. Felker, H. Ng, C. Jin, Red Opal: product-feature scoring from reviews, in *8th ACM Conf. on Electronic Commerce (EC)*, no. 10, pp. 182–191, San

- Diego, California (2007). <https://doi.org/10.1145/1250910.1250938>. <https://doi.org/10.1145/1250910.1250938>
11. H. Nakagawa, T. Mori, A simple but powerful automatic term extraction method, in *Int. Workshop on Computational Terminology*, Morristown, NJ, USA (2002)
 12. C. Cardie, J. Wiebe, T. Wilson, Litman, Combining low-level and summary representations of opinions for multi-perspective question answering, new directions in question answering, in *Papers from 2003 AAAI Spring Symposium*, Stanford University, Stanford, CA, USA, January 2003
 13. B. Singh, S. Kushwah, S. Das, Multi-feature segmentation and cluster-based approach for product feature categorization, *International Journal of Information Technology and Computer Science (IJITCS)*, MECS Publisher. IJITCS **8**(3), 33–42 (2016)
 14. B. Singh, S. Kushwah, S. Das, P. Johri, Issue and challenges of online user generated reviews across social media and E-commerce website, in *Proceeding of IEEE International Conference on Computing Communication and Automation (ICCCA-2015)*, 15–16 May, pp. 818–822 (2015). <https://doi.org/10.1109/CCAA.2015.7148486>
 15. S. Das, B. Singh, S. Kushwah, P. Johri, Opinion based on polarity and clustering for product feature extraction. *Int. J. Information Eng. Electronic Bus. (IJIEEB)* **8**(5), 33–42 (MECS Publisher, August 2016. ISSN: 2074-9023 (Print), ISSN: 2074-9031, Online). <https://doi.org/10.5815/ijieeb2016>.
 16. S. Kushwah, S. Das, Sentiment analysis of big-data in healthcare: issue and challenges, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India (2020), pp. 658–663. <https://doi.org/10.1109/ICCCA49541.2020.9250841>
 17. S. Kushwaha, S. Das, Hierarchical agglomerative clustering approach for automated attribute classification of the health care domain from user generated reviews on web 2.0, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India (2020), pp. 671–676. <https://doi.org/10.1109/GUCON48875.2020.9231122>
 18. S. Homoceanu, M. Loster, C. Lofi, W. Balke, Will i like it? Providing product overviews based on opinion excerpts, in *2011 IEEE 13th Conference on Commerce and Enterprise Computing*, Luxembourg (2011), pp. 26–33. <https://doi.org/10.1109/CEC.2011.12>
 19. <https://aliasi.com/lingpipe/demos/tutorial/sentiment/readme.html>. LingPipe—tool kit for processing text using computationallinguistics
 20. Esuli, F. Sebastiani, SENTIWORDNET: a publicly available lexical resource for opinion mining, in *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, Genoa, Italy, May 2006 European Language Resources Association (ELRA). https://www.lrec-conf.org/proceedings/lrec2006/pdf/384_pdf.pdf
 21. S. Baccianella, A. Esuli, F. Sebastiani, SENTIWORD NET 3.0: an enhanced lexical resource for sentiment analysis and opinion mining, in *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May 2010, European Language Resources Association (ELRA). <https://www.lrec-conf.org/proceedings/lrec2010/pdf/769>
 22. <https://www.amazon.com/>. Online Shopping for Customers
 23. <https://nlp.stanford.edu/software/tagger.shtml>. The Stanford Natural Language Processing Group
 24. E. Riloff, J. Wiebe, Learning extraction patterns for subjective expressions, in *Association for Computational Linguistics*, No. 8, pp. 105–112, EMNLP '03, USA (2003). <https://doi.org/10.3115/1119355.1119369>
 25. M. Gamon, A. Aue, S. Corston-Oliver, E.K. Ringger, Pulse: mining customer opinions from free text, in *Advances in Intelligent Data Analysis VI. IDA 2005* ed. by A.F. Famili, J.N. Kok, J.M. Peña, A. Siebes, A. Feelders. Lecture Notes in Computer Science, vol 3646 (Springer, Heidelberg, 2005). https://doi.org/10.1007/11552253_12
 26. K. Kageura, B. Umino, Methods of automatic term recognition: a review. *Terminology* **3**, 259–289 (1996)

27. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering* ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh. Lecture Notes in Electrical Engineering, vol 661 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-15-4692-1_58
28. A.M. Popescu, O. Etzioni, Extracting product features and opinions from reviews, in *Natural Language Processing and Text Mining* ed. by A. Kao, S.R. Potet (Springer, London, 2007). https://doi.org/10.1007/978-1-84628-754-1_2
29. Z. Zhai, B. Liu, H. Xu, J. Peifa, Clustering product features for opinion mining, in *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, WSDM '11*, no. 8, pp. 347–354. ACM, Hong Kong, China (2011). <https://doi.org/10.1145/1935826.1935884>
<https://doi.org/10.1145/1935826.1935884>
30. V. Hatzivassiloglou, W.M. Janyce, Effects of adjective orientation and gradability on sentence subjectivity, in *Proceedings of the 18th Conference on Computational Linguistics*, vol. 1, no. 7, pp. 299–305, Saarbrücken, Germany, COLING '00 (2000). Association for Computational Linguistics. <https://doi.org/10.3115/990820.990864>. <https://doi.org/10.3115/990820.990864>

A Mutual Authentication and Key Agreement Protocol for Vehicle to Grid Technology



Azmi Iqbal, Akber Ali Khan, Vinod Kumar, and Musheer Ahmad

Abstract Vehicle-to-grid (V2G) technology expedites the electric vehicles to distribute additional electricity into the grid systems as well as it helps to get back from the power grid through charging. An efficient key framework is very important for the initiation of a bi-directional supply of electricity into and out of the device. In order to effectively introduce V2G communication, the key agreement and authentication framework must be spared from cybersecurity attacks. Safe and effective key settings are surely an essential issue for Internet-based smart grid system. An authentication protocol gives secure communication among clients and specialist organizations for security and assurance reasons. A couple of authentication protocols are open in the literature. Be that as it might, they are engaged to acknowledge security assaults effectively, or on the other hand, they are not computationally efficient. In our protocol, we structure an ECC-based regular confirmation protocol for vehicle to grid communication utilizing password change approach. The proposed protocol secure against many security attributes. In addition, the proposed framework takes considerably less communication and computational costs than other V2G implementations, which are beneficial for practice.

Keywords Mutual authentication · Elliptic curve cryptography · Security and privacy · Vehicle to grid

A. Iqbal · A. A. Khan (✉) · M. Ahmad
Department of Applied Sciences and Humanities, Jamia Millia Islamia, New Delhi, India
e-mail: cs.akberkhan@gmail.com

A. Iqbal
e-mail: azmirahman95@gmail.com

M. Ahmad
e-mail: mahmad@jmi.ac.in

V. Kumar
Department of Mathematics PGDAV College, University of Delhi, New Delhi, India
e-mail: vinod@pgdav.du.ac.in; vinodiitkgp13@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_66

1 Introduction

The prototype of the combination of advanced power and electronics technology which refers to a single term called Energy Internet model(EI model), aims at resemblance of various technologies of Information and Communications technology methods, different contemporary computerized networks and the electric-powered capacity of the present-day systems that could be helpful in developing smart next-generation systems tremendous grids [1]. The idea of EI makes it possible for energy to similar data being exchanged on the conventional Internet where it can be shared. This EI model decides to incorporate the various aspects of Economics, knowledge on data and energy that depends on power-grid build network, which finally results in an unfolded structure for sharing power grid-based network, which finally results in an open structure for sharing energy and the corresponding data. It allows the integration of different SG energy sources and leads to a comprehensive productivity in energy generation, delivery, customers energy generation, intended operations and provision of services. Both of these bestowing variables are focused on the safe contact between the participating organizations. The technology of the V2G makes bidirectional movements of energies between the power grid and the electric grid control electric vehicles. It is possible to transfer electricity whenever the later needs to get recharged from the power grid to the EV and in case of extra power it's got battery from the EV to power grid produced. Therefore, as a result, in the V2G system, vehicles can operate as an energy forte for the smart-grid and hence, this type of bidirectional charging will make a major contribution to the generation of energy. Due to this V2G technique, an individual owner or a single household of the Energy Internet (EI) can engage in the sale and purchase of energy from an EV without even a traditional power generation and its distribution system. ISO/IEC/IEEE 18880 [2] is the standard dealing with EI-based different protocols and architectures. For different protocols, this specification specifies the data storage, architecture and application services. In the subsequent standards such as ISO/IEC/IEEE 18881-3, the security weaknesses have been resolved and also the functionality has been improved. The V2G network framework focuses on merely charging a vehicle, a grid-based charging station and a grid power supply system. Hence, subsequently many users can share their charging stations at the same time. Different grid protocols are used at different charging speeds. The IEC 15118 standard is used for the charging purposes to create contact between electric vehicles [EV] and electric vehicles charger [3] protocol. The EV delivery services also use the Open Charge Point Protocol (OCPP) to communicate with the energy control systems given in [4]. The protection of the Energy Internet-based Vehicle-to-grid systems is one of the crucial distresses as the vehicle may get targeted by any attacker that effects not only its operations but also its privacy. This privacy concern may be the revealing off the identity of the client, the location and the driving path of the vehicle etc. As a part of the vehicle mobility, safety protocols are becoming increasingly difficult to enforce in the V2G infrastructure-based world. A few other researches, recently, can be seen to ensure

generally positive key agreement for Energy Internet-based vehicle-to-grid systems [1, 2].

1.1 Related Work

To establish secure data sharing all the times, secure data communication is one of the most analytical basis for the energy-Internet framework. To ensure a safe and efficient exchange of data between components, high security and performance protocols are required. To deal with this issue, A number of researchers have suggested a number of shared authenticated protocols and key establishment schemes, which is best suited for the Advanced Metering Infrastructure (AMI) with different safety requirements and objectives. Two major identity-based approaches for the establishment of the ECC are suggested by Mohammadali et al. [5]. Protocols minimize the computational overhead of the SM side of the AMI, they are immune to replay and attacks on desynchronizing. Be that as it may, they are susceptible to impersonation assaults, man-in-the-middle assaults, reply assaults and incur high computational costs during key times establishing. Two key interchange protocols, based on the use of symmetric-key algorithm and elliptic curve cryptography, implemented by Nicanfar et al. [6]. Wu and Zhou [7] put forward an authentication and key distribution protocol by enhancing the addition of two major keys of cryptography scheme such as public key and symmetric key, and on this study, they proclaim that these schemes will debar reply and man-in-the-middle assaults which usually happens. On the other hand, Xia and Wang [8] have proved with their theory that Nicanfar et al. [6] cannot promise protection against man-in-the-middle attacks and further they proposed a new data aggregation scheme [8]. However, Park et al. [9] verified that the scheme presented by Xia and Wang [8] is not safe against on impersonation assaults [9]. Furthermore, the privacy conditions of the consumer cannot be addressed. Tsai et al. [10] setup an identity-based signature scheme and ID encryption scheme [10] for smart grid key distribution. Odelu et al. [11] researched and demonstrated the protocol outlined [10], that it cannot guarantee the privacy of the smart meters and the security of the session key [11]. They argued a new protocol that could reduce overheads for computing. However, Chen et al. demonstrated that the protocol introduced in [11] is powerless against a few assaults, and it has huge computational and communication costs. In addition to this, the experiment reveals that the protocol in [11] become ineffective against the man-in-the-middle assaults which may sometimes produce Dos attack at the server end. In [12], the researchers considered the physical protection and suggested a protocol using the idea of physical unlovable functions. Few privacy concerns in V2G communications have appeared in literature [13–18] and [19]. In these protocols, protection of the vehicle proprietor is considered as a significant concern. Additionally, the vast majority of these protocols cannot grantee the area security of the electric vehicle user, which is fundamental for safely observing the status of the EV and efficiently offering types of assistance of the EV user.

1.2 Contribution

The proposed paper has the following features:

- The proposed protocol has key agreement between particular vehicle user and grid server.
- The proposed protocol maintains many cryptographic properties and secure against many attacks in authentication protocols.
- The proposed protocol is much appropriate for practical applications in V2G communication as it provided extra security features along with the lower communication and computation costs.

1.3 Organization of the Paper

The rest of paper as follows:

Section 2, gives mathematical preliminaries. Section 3, the proposed protocol. Section 4, security analysis. Further, performance analysis are given in Sect. 5. Finally, we conclude the results in last Section.

2 Preliminaries

Preliminaries given in this section.

2.1 Notations

Following useful notations throughout the paper are listed in below (Table 1).

2.2 Elliptic Curve Cryptography

E be an elliptic curve over a real prime finite field describe as $E_p(a, b) : v^2 = u^3 + au + b \pmod{q}$, where $a, b \in \mathbb{Z}_p^*$ and $4a^3 + 27b^2 \not\equiv 0 \pmod{q}$ non singular elliptic curve mod $q \neq 0$, defined as $G = \{(u, v) : u, v \in \mathbb{Z}_p^*, (u, v) \in E\} \cup \{\Theta\}$, where the point Θ known as identity of G under addition.

Table 1 Notations

Notations	Description
G	Elliptic curve group under addition
E	Elliptic curve
$V2G$	Vehicle to grid
ID_V	The unique Identity of vehicle V_i
ID_S	The unique Identity of server S_i
p and q	Large prime numbers
$E_p(\cdot)$	Elliptic curve over a real prime finite field
Z_q^*	Finite field of integer of order $q - 1$
SG	Smart grid
g	The base point on G
V	Vehicle user
SK_{ij}	Session key agreement between participants i and j
\parallel	The concatenation operator
$h(\cdot)$	Hash function
\oplus	Bitwise XOR operation
\mathcal{A}	An Adversary
$\stackrel{?}{=}$	Whether equal or not

1. If $M = (u_1, v_1) \in G$ and $N = (u_2, v_2) \in G$, then $M + N = (u_3, v_3)$, where $u_3 = \lambda^2 - u_1 - u_2 \pmod p$, $v_3 = (\lambda(u_1 - u_3) - v_1) \pmod q$ and

$$\lambda = \begin{cases} \frac{v_2 - v_1}{u_2 - u_1} \pmod q & \text{if } M \neq N \\ \frac{3u_1^2 + a}{2v_1} \pmod q & \text{if } M = N \end{cases}$$

2.The scalar multiplication on G defined as $kM = M + M + M + M + \dots$ M ($k - \text{times}$). Additional information are given in [19–22].

3 The Proposed Scheme

This section contains following phases:

3.1 Initialization Phase

In initialization phase

- Step 1. Selects $p, E_p(a, b) : v^2 = u^3 + au + b \pmod p$. where $P \in E(F_p), a, b \in F_p$ with $4a^3 + 27b^2 \pmod p \neq 0$.
- Step 2. Chooses $h(\cdot)$.
- Step 3. Generates its private key as $s \in Z_q^*$ and public key as $P_{pub} = sP$.
- Step 4. Disclose $\{E_p(a, b), p, q, Pub, h(\cdot)\}$ and keep s secretly.

3.2 Registration Phase

In this phase, V and SG registration as follows.

- Step 1. To register with smart grid server, V selects his or her ID_V and generates $v \in Z_q^*$, forwards $\{ID_V \oplus v, ID_V\}$ toward smart grid server through secure medium.
- Step 2. On obtaining $\{ID_V \oplus v, ID_V\}$. Verify ID_V . If it is not exist then provide password PW_V and counter C_V . If verification is successfully then server computes $HS_1 = h(ID_V \oplus v || s || C_V)$ where, s is the private key of smart grid server and C_V is counter. Further, server computes $HS_2 = HS_1 \oplus h(ID_V \oplus PW_V)$ and store $\{HS_2, C_V, PW_V\}$ in database and send to entity V .
- Step 3. On getting $\{HS_2, C_V, PW_V\}$ after that, vehicle compute $HS = HS_2 \oplus h(ID_V || PW_V)$, compute $PW_V = h(PW_V || HS || ID_V)$ and store $\{PW_V, C_V\}$ in database of V (Table 2).

Table 2 Registration of vehicle

V	SG
Inputs ID_V Selects $v \in Z_q^*$ Sends $\{ID_V \oplus v, ID_V\}$ \implies	Verify ID_V in database. If it is not exist then provide password PW_V and counter C_V . Computes $HS_1 = h(ID_V \oplus v s C_V)$ Computes $HS_2 = HS_1 \oplus h(ID_V \oplus PW_V)$ Sends $\{HS_2, C_V, PW_V\}$ \longleftarrow
Computes $HS = HS_2 \oplus h(ID_V PW_V)$ Computes $PW_V = h(PW_V HS ID_V)$ Store $\{PW_V, C_V\}$ in its database	

3.3 Login and Authentication Phase

The explanation of authentication phase are also shown in Table 3 and given as follows:

- Step 1. Vehicle user login with his/her ID_V^* and password PW_V' . Further, vehicle user computes $PW_V' = h(PW_V' \| HS \| ID_V')$ and verifies $PW_V' \stackrel{?}{=} PW_V$. If yes then, Generates $x \in Z_q^*$, Computes $A=x.g$, $H_1 = h(PW_V \| C_V \| ID_V)$, $ID_{V1} = ID_{V1} \oplus h(HS \| C_V)$ and send $M_1 = \{A, H_1, ID_{V1}, t_1\}$ toward through a reliable medium.
- Step 2. On getting M_1 , smart grid server verifies $t_4 - t_3 \leq \Delta t$, if verification is successfully done then, smart grid compute $ID_V^* = ID_{V1} \oplus h(HS_1 \| C_V)$, $H_1^* \stackrel{?}{=} H_1$. After that, generates a random number $y \in Z_q^*$, computes $B = y.g$, computes $H_2 = h(ID_V^* \| ID_S \| A \| H_1^* \| t_1)$ computes session key $SK_S = h(H_2 \| ID_V^* \| ID_S \| y.A \| t_3)$ and computes $ID_{S1} = ID_S \oplus h(H_1^* \| C_V \| A)$ further, smart grid server $M_2 = \{ID_{S1}, B, H, t_3\}$.
- Step 3. On receiving M_2 , vehicle verifies $t_4 - t_3 \leq \Delta t$ if yes then, V computes $ID_S^* = ID_{S1} \oplus h(H_1 \| C_V \| A)$, computes $H_2^* = h(ID_V \| ID_S^* \| A \| H_1 \| t_1)$ and verifies $H_2^* \stackrel{?}{=} H_2$, V computes $SK_V = h(H_2^* \| ID_V \| ID_S^* \| x.B \| t_3)$ Thus session key $SK_V = SK_S = SK$.

3.4 Password Change Phase

- Step 1. Input ID_V and PW_V then computes $PW_V' = h(PW_V \| HS \| ID_V')$.
- Step 2. V verifies $PW_V' \stackrel{?}{=} PW_V$ if not terminate the process otherwise select new identity ID_V^{NEW} and password PW_V^{NEW} and computes $PW_V^{NEW} = h(PW_V^{NEW} \| HS \| ID_V^{NEW})$.
- Step 3. V replaces PW_V by PW_V^{NEW} and ID_V by ID_V^{NEW} .

4 Security Analysis

Here, we describe the analysis of proposed protocol. We first demonstrate that our proposed scheme is secure.

Table 3 Authentication phase

V	SG
Inputs ID'_V and PW'_V Computes $PW'_V = h(PW'_V \ HS \ ID'_V)$ Verifies $PW'_V \stackrel{?}{=} PW_V$ Generates $x \in Z_q^*$ Computes $A = x.g$ Computes $H_1 = h(PW_V \ C_V \ ID_V)$ Computes $ID_{V1} = ID_V \oplus h(HS \ C_V)$ Sends $\{A, H_1, ID_{V1}, t1\}$ →	Verifies $t_2 - t_1 \leq \Delta t$ Computes $ID_V^* = ID_{V1} \oplus h(HS_1 \ C_V)$ Computes $H_1^* \stackrel{?}{=} H_1$ Generates $y \in Z_q^*$ Computes $B = y.g$ Computes $H_2 = h(ID_V^* \ ID_S \ A \ H_1^* \ t_1)$ Computes $SK_S = h(H_2 \ ID_V^* \ ID_S \ y.A \ t_3)$ Computes $ID_{S1} = ID_S \oplus h(H_1^* \ C_V \ A)$ Sends $M_2 = \{ID_{S1}, B, H, t_3\}$ ←
Verifies $t_3 - t_4 \leq \Delta t$ Computes $ID_S^* = ID_{S1} \oplus h(H_1 \ C_V \ A)$ Computes $H_2^* = h(ID_V \ ID_S^* \ A \ H_1 \ t_1)$ Verifies $H_2^* \stackrel{?}{=} H_2$ Computes $SK_V = h(H_2^* \ ID_V \ ID_S^* \ x.B \ t_3)$ hence session key $SK_V = SK_S = SK$	

4.1 Replay Attack

Any session is refreshed by the time stamps $t_j - t_i \leq \Delta t$, where ΔT is the valid time period which are verified respectively by vehicle and server. In addition, vehicle and server generate random numbers $x \in Z_q^*$ and $y \in Z_q^*$. Hence, the adversary cannot replay a message from a session.

4.2 User Anonymity

V send its $ID_{V1} = ID_V \oplus h(HS \| C_V)$ to the smart grid then, smart grid compute anonymous identity $ID_V^* = ID_{V1} \oplus h(HS_1 \| C_V)$ of V. SG sends its partial identity

$ID_{S1} = ID_S \oplus h(H_1^* \| C_V \| A)$ to the vehicle V then, V compute anonymous identity $ID_{S1} = \oplus h(H_1 \| C_V \| A)$. Hence, the proposed scheme maintain said property.

4.3 Message Authentication

Smart grid receives the message $M_1 = \{A, H_1, ID_{V1}, t_1\}$. Now check the authenticity by validating the time-stamp condition $t_2 - t_1 \leq \Delta t$ and checks the validity by hash function $H_1^* \stackrel{?}{=} H_1$. Vehicle gets the message $M_2 = \{ID_{S1}, B, H_2, t_3\}$. Verifies the authenticity by validating $t_4 - t_3 \leq \Delta t$ and $H_2^* \stackrel{?}{=} H_2$. Thus message acquired within verifying conditions and hash functions, which is hard to obtained for \mathcal{A} . Therefore, scheme withstand the message authenticity.

4.4 Key Freshness Property

In this proposed protocol we used fresh random number and fresh time stamp at each step of our protocol to maintain keep freshness property of cryptography.

4.5 Man-in-the-Middle Attack

- Step 1. \mathcal{A} may used the past information to enter in the server. \mathcal{A} replay $M_1 = \{A, H_1, ID_{V1}, t_1\}$ where $H_1 = h(PW_V \| C_V \| ID_V)$, fresh number $x \in Z_q^*$.
- Step 2. Upon obtain M_1, H_1 or verifies $t_i - t_j \leq \Delta t$ and $H_i^* \stackrel{?}{=} H_i$. The private keys of smart grid and vehicle cannot be accessed by \mathcal{A} . \mathcal{A} is not able to calculate a real verifier. Therefore, \mathcal{A} cannot adjust a parameter as a result of the verifiers should be changed appropriately. Hence, the proposed scheme withstands with man-in-the-middle assaults.

4.6 Security Against Impersonation Attack

Any attacker \mathcal{A} can try as a V to login the SG server. \mathcal{A} get the information $M_1 = \{A, H_1, ID_{V1}, t_1\}$ and try to compute $\{A, H_1\}$ where $A = x.g$ is an elliptic curve point which is not easy for any adversary to compute A and H_1 is the hash value which contain password and counter of vehicle user which is not easy to guess is our proposed protocol. Hence, any \mathcal{A} cannot impersonate message M_1 similarly any adversary cannot impersonate server sight. Hence, the proposed protocol maintain impersonation from any adversary.

4.7 Session Key Agreement

In the proposed protocol, vehicle and smart grid both computes $SK_V = h(H_2^* || ID_V || ID_S^* || x.B || t_3)$ and $SK_S = h(H_2 || ID_V^* || y.A || t_3)$, respectively. It is clear that $SK_V = SK_S$. Hence, the communication between them is secure.

5 Performance Analysis

We explain the detail description of performance analysis with related schemes [5–7, 10, 11]. That are completed in three stages as follows:

5.1 Security-Based Comparison

In Table 4, we draw the comparison with [5–7, 10, 11].

5.2 Computation Cost Comparison

In section, we compute the computational overhead of the proposed scheme and compare with related framework. The operations used in computation cost performed on HTC One smartphone cortex A9, MPcore processor operating at 890 MHz and a personal computer, virtual machine Core i5 4300, dual core 2.60GHz processor using Ubuntu 12.04 operating system [12].

Table 4 Security features comparison

Scheme	R2	R1	R4	R3	R5	R6
[5]	Yes	Yes	No	No	Yes	No
[6]	No	No	No	No	No	No
[7]	No	No	No	No	No	No
[10]	Yes	Yes	Yes	Yes	No	No
[11]	Yes	Yes	Yes	Yes	Yes	No
Proposed	Yes	Yes	Yes	Yes	Yes	Yes

R2: Customer privacy; **R1:** Protection against eavesdropper; **R4:** Resist against man-in-the-middle assaults; **R3:** Forward secrecy; **R5:** Session key security; **R6:** Resist against DoS assaults

Table 5 Execution time of various cryptographic operations

Operation	V (ms)	SG (ms)
Multiplication operation	21.86	14.5
Hash operation	0.0186	0.011
Modular exponential operation	7.23	2.34
Symmetric encryption/decryption operation	0.0584	0.41
Bilinear pairing	8.67	3.78
Multiplication point operation	5.12	2.6
Certificate generation/Verification operation	55.946	2.6

Table 6 Computation cost comparison

Scheme	Users device	CS	Total cost (ms)
[5]	$3T_{MP} + T_M + T_{CERT_{gen}} + 3T_H \approx 88.15$	$3T_{MP} + T_M + T_{CERT_{ver}} + 4T_H \approx 57.87$	146.02
[10]	$4T_{MP} + T_E + 5T_H \approx 27.85$	$4T_{MP} + T_E + 5T_H \approx 23.22$	51.37
[6]	$3T_{MP} + T_M + T_{CERT_{gen}} + T_H \approx 93.24$	$4T_{MP} + T_M + T_{CERT_{ver}} + 4T_H + T_S \approx 63.77$	157.01
[7]	$2T_{MP} + T_M + T_{CERT_{gen}} + T_H + T_S \approx 92.38$	$3T_{MP} + T_M + T_{CERT_{ver}} + 3T_H + T_S \approx 57.88$	150.26
[11]	$3T_{MP} + T_E + 6T_H \approx 22.74$	$2T_{MP} + T_E + 2T_B + 6T_H \approx 15.32$	38.06
Proposed	$T_{MP} + 3T_H \approx 5.1758$	$T_{MP} + 2T_H \approx 2.622$	7.7987

T_M : Multiplication operation; T_S : Symmetric encryption/decryption operation; T_E : Modular exponential operation; T_{MP} : Multiplication point operation; T_B : Bilinear pairing; T_H : Hash operation; $T_{CERT_{gen/ver}}$: Certificate generation/Verification operation (Tables 5 and 6).

5.3 Communication Cost Comparison

Table 7 shows the communication comparison cost with related frameworks [5–7, 10, 11]. We have taken time stamp to be 32, 160 for random number, 256 for symmetric encryption/decryption, 60 for identifier, output of hash function 160, bilinear pairing takes and elliptic curve point 320 and digital signature takes 1024 are in bits [12].

Table 7 Communication cost comparison

Scheme	Number of messages	Communication costs in bits
[5]	4	2340
[10]	4	6880
[6]	4	2176
[7]	5	4064
[11]	7	2912
Proposed	2	1152

6 Conclusion

Vehicle to grid is a framework for our increasingly linked Internet society and ensure reliability and stability. Vehicle to grid protection is a topic of greater concern to government in business, academia and so on. To guarantee secure and efficient communication between vehicle user and grid server, we presented an ECC-based key agreement and authentication framework for V2G environment. We have shown that our protocol manages different security attacks. We have also compared the cryptographic security attacks and with other protocols. In addition, the proposed protocol has far fewer cost of computational and communication as compared with other protocols already in existence. Our results confirm that the proposed scheme is efficient for interacting with V2G.

References

1. Z. Dong, *Towards an Intelligent Future Energy Grid*, The University of Sydney, New South Wales
2. V.C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, G.P. Hancke, A survey on smart grid potential applications and communication requirements. *IEEE Trans. Industrial Informatics* **9**(1), 28–42 (2012)
3. S. Lee, Y. Park, H. Lim, T. Shon, Study on analysis of security vulnerabilities and countermeasures in iso/iec 15118 based electric vehicle charging technology, in *International Conference on IT Convergence and Security (ICITCS)*, IEEE, 1–4 (2014)
4. C. Alcaraz, J. Lopez, S. Wolthusen, Ocpp protocol: security threats and challenges. *IEEE Trans. Smart Grid* **8**(5), 2452–2459 (2017)
5. A. Mohammadali, M.S. Haghghi, M.H. Tadayon, A. Mohammadi-Nodooshan, A novel identity-based key establishment method for advanced metering infrastructure in smart grid. *IEEE Trans. Smart Grid* **9**(4), 2834–2842 (2016)
6. H. Nicanfar, V.C. Leung, Multilayer consensus ecc-based password authenticated key-exchange (mcepak) protocol for smart grid system. *IEEE Trans. Smart Grid* **4**(1), 253–264 (2013)
7. D. Wu, C. Zhou, Fault-tolerant and scalable key management for smart grid. *IEEE Trans. Smart Grid* **2**(2), 375–381 (2011)
8. J. Xia, Y. Wang, Secure key distribution for the smart grid. *IEEE Trans. Smart Grid* **3**(3), 1437–1443 (2012)

9. J.H. Park, M. Kim, D. Kwon, Security weakness in the smart grid key distribution scheme proposed by Xia and Wang. *IEEE Trans. Smart Grid* **4**(3), 1613–1614 (2013)
10. J.-L. Tsai, N.-W. Lo, Secure anonymous key distribution scheme for smart grid. *IEEE Trans. Smart Grid* **7**(2), 906–914 (2015)
11. V. Odelu, A.K. Das, M. Wazid, M. Conti, Provably secure authenticated key agreement scheme for smart grid. *IEEE Trans. Smart Grid* **9**(3), 1900–1910 (2016)
12. P. Gope, B. Sikdar, An efficient privacy-preserving authentication scheme for energy internet-based vehicle-to-grid communication. *IEEE Trans. Smart Grid* **10**(6), 6607–6618 (2019)
13. D.P. Ghosh, R.J. Thomas, S.B. Wicker, A privacy-aware design for the vehicle-to-grid framework, in *2013 46th Hawaii International Conference on System Sciences*, IEEE, pp. 2283–2291 (2013)
14. H. Nicanfar, P. TalebiFard, S. Hosseinienezhad, V.C. Leung, M. Damm, Security and privacy of electric vehicles in the smart grid context: problem and solution, in *Proceedings of the Third ACM International Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications*, pp. 45–54 (2013)
15. H. Liu, H. Ning, Y. Zhang, L.T. Yang, Aggregated-proofs based privacy-preserving authentication for v2g networks in the smart grid. *IEEE Trans. Smart Grid* **3**(4), 1722–1733 (2012)
16. A. Abdallah, X.S. Shen, Lightweight authentication and privacy-preserving scheme for v2g connections. *IEEE Trans. Vehicular Technol.* **66**(3), 2615–2629 (2016)
17. N. Saxena, B.J. Choi, Authentication scheme for flexible charging and discharging of mobile vehicles in the v2g networks. *IEEE Trans. Information Forensics Security* **11**(7), 1438–1452 (2016)
18. D. He, S. Chan, M. Guizani, Privacy-friendly and efficient secure communication framework for v2g networks. *IET Commun.* **12**(3), 304–309 (2017)
19. A.A. Khan, V. Kumar, M. Ahmad, S. Rana, D. Mishra, Palk: password-based anonymous lightweight key agreement framework for smart grid. *Int. J. Electrical Power Energy Syst.* **121**, 106121 (2020)
20. V. Kumar, S. Jangirala, M. Ahmad, An efficient mutual authentication framework for healthcare system in cloud computing. *J. Med. Syst.* **42**(8), 142 (2018)
21. A.A. Khan, V. Kumar, M. Ahmad, An elliptic curve cryptography based mutual authentication scheme for smart grid communications using biometric approach. *J. King Saud Univ.-Comput. Information Sci.* (2019). <https://doi.org/10.1016/j.jksuci.2019.04.013>
22. V. Kumar, A.A. Khan, M. Ahmad, Design flaws and cryptanalysis of elliptic curve cryptography-based lightweight authentication scheme for smart grid communication, in *Advances in Data Sciences, Security and Applications*. Springer, pp. 169–179 (2020)

Fire Detection and Real Time Monitoring Systems Through IoT Sensors



Sudip Suklabaidya and Indrani Das

Abstract The Internet of Things (IoT) has become the new area of research and development for variety of applications in industrial and domestic areas. The model proposed in the paper consists of a built-in Wifi module Node MCU, an open source IoT platform, Arduino UNO controller, and programming languages along with sensor devices and network. The proposed model translates, verifies data and communicates to the server based on working principles and framework. It also develops IPv6 network environment which is capable of receiving and responding to different formats of data received from the sensors given in different formats and assess their relevance and accuracy with the help of appropriate coding and forward them to the server in a graphical format. It is assumed that the system will be used for highly emergency situation to make decisions regarding response to the fire in adverse state. The experimental results show the effectiveness and reliability of our proposed system and are illustrated with the graph where data are inserted into the ThingSpeak.com server.

Keywords IoT · Node-MCU · Arduino UNO · Sensor · ThingSpeak IoT Cloud

1 Introduction

The Internet of Things (IoT) has revolutionized human living. It spans between virtual and physical domains and make use of tools such as simulators and test-beds to frame the initial proof-of-concepts and subsequent prototypes [1]. IoT is a physical network of objects where various sensors are being monitored and controlled remotely using existing network infrastructure. It can incorporate several techniques and mechanism

S. Suklabaidya (✉)

Department of Computer Science and Application, Karimganj College, Assam, India

I. Das

Department of Computer Science, Assam University, Silchar, India

for better accuracy and operation. The sensor can sense rise in temperature, flame and send these data to a monitoring station to generate needful instructions or execute appropriate decision [2].

Since millions of smart devices are connected to the internet and they have to be communicate through a unique address space, addressing and identification is a major weaknesses in IoT. Also communication of data between smart objects especially through wireless mode is a power consuming task as IoT devices are battery powered having minimum computation capability and storage capacity. These leads to various constraints and issues on IP enabled wireless sensor network. To overcome the current weaknesses and fragmentation of the internet of things, the usage of IPv6 and other standards protocols were used [3]. The main challenges are:

- (a) Investigate the strength of IPv6 extended features and interrelated standards for analyzing the facts of IoT and its overheads.
- (b) Build a prototype to provide a mechanism for analysis, search for an integration of services with the help of IoT enabled keen objects to connect clients, servers and various other devices through IPv6 [3].
- (c) To parallel the device mobility along with mobile phone network integration based on the programming methodology and the prototype used.

The Internet protocol is a sole requirement for any connection over internet. The limited size and space of IPv4 has led to the transition to IPv6. As IPv6 provides a highly scalable address scheme of 2^{128} unique addresses over IPv4 2^{32} unique addresses. IPv6 also provides end to end connectivity in a secured and dispersed routing mechanism.

Now-a-days homes and commercial buildings are equipped with various kinds of sensors and controlled by many latest technologies. In addition to fire detectors, sensors are fixed dedicated to monitoring its internal environment, security and integrity. Apart from these they monitored factors like smoke, temperature, acceleration, gases, humidity, pressure, security images (CCTV), intrusion etc. The output from each of these sensors is capable of providing system with valuable data for use in forecasting the progression and the response of the building and its occupants to the fire. The sensors provide only a simple digital (on/off) signal, while others may provide a -5 V to $+5\text{ V}$ equivalent signal or a signal of a few milli volts.

Fire response System is increasingly providing design solutions for such homes and commercial buildings that facilitate function while providing a first line of defense against the occurrence and propagation of fire. This system aims to give people a low cost and efficient way to fireproof their homes, offices, storage units. The temperature module also allows users to monitor the temperature and humidity thus never allowing users to compromise their cold storages, green house etc. and all such infrastructure which require a certain range of temperature to be maintained. The components required are very much affordable and readily available in the market. Moreover the installation procedure is quite easy which requires programming knowledge in NodeMCU and Arduino as if in some case the user requires customization. A strong network must be in place to collect data from the various sensors. It is also anticipated that sensors can continuously send data to the server for

as long as possible during a fire, though it is expected that in the course of a significant fire many sensors, and the sections of network to which they are connected, will be demolished.

In this paper we have developed a model fire detection system using sensor like LM35 (Temperature Sensor), MQ-2(Smoke Sensor), YL-44(Buzzer Module) and Flame sensor. All these sensors along with WiFi module Node-MCU are connected through Arduino Board, an 8-bit microcontroller which generates an Integrated control System. The system on receiving input are analyzed and produces real time output through Thingspeak Server in graphical format. The main objective of this paper is to build a system through which information can be send remotely to the user or to the competent authority when the fire is being detected.

The paper is synchronized as follows: Sect. 2 in the paper highlights review of the previous related works done to solve the stated problem, Sect. 3 describe the proposed approach with flowchart, Sect. 4 elaborates the experimental setup and working mechanism of the proposed system. The Sect. 5 present the results and findings and finally Sect. 6 concludes the paper.

2 Related Work

In [4] the author have proposed an IoT based Theft Pre-emption and Security System. The security system proposed here is based on IoT, in order to prevent risk and theft in home, bank etc. The paper also highlights to reduce human work through automation as it is an important aspect for security system.

Multi-sensor and information fusion technology based on Dempster-Shafer evidence theory is applied in the system of a building fire alarm to realize early detecting and alarming [5]. The paper uses various sensors to examine the parameters of the fire process in space and time which are then expanded and compared with a single-sensor system.

IoT-based Intelligent for Fire Emergency Response Systems [6, 7].The paper proposes to reduce casualties by determining the point of occurrence of a disaster in a building to prevent directional confusion of the emergency lights and inappropriate evacuation guidance.

Here in this paper “Microcontroller ATmega 328P and GSM Based Advanced Home Security System” [8], the author provides higher security to homes. They have used both fingerprint and password for authentication and hence reduce the probability of hacking.

In [9] they proposes a model that include PIR sensor, temperature sensor, heat sensor and gas sensor. On top of these sensors, the footage from the camera is also used to detect the fire through image processing. The main advantage of this system is that it has a very high accuracy. If the fire has been detected a mail is sent to the security and the nearest fire department with an attachment of the photo.

The Paper “Development of Fire Alarm System using Raspberry Pi and Arduino Uno” [10] proposes a fire alert system in a real-time monitoring system that detects

the presence of smoke in the air due to fire and captures images via a camera installed inside a room when a fire occurs. The system will display an image of the room when the smoke is detected.

In this paper [11], Remote Household Appliance Control System Using GSM, the author build a system through GSM technology. The system will facilitate the user or the owner to monitor and control his household appliances through handheld set by sending messages.

3 Proposed Model and Architecture

3.1 Working Mechanism

The working methodology of the proposed fire detection system has been depicted in the Fig. 1. The figure shows that how the components are integrated in a breadboard. The temperature sensor LM 35 along with smoke sensor MQ-2 and flame sensor used in the system are connected to a WiFi module Node MCU and the buzzer sensor is connected with the Arduino UNO board. Having being connected in their respective places the Arduino UNO now loads its library which is shown with the sketch option inside the IDE and is regulated with all library for proper functioning. Parallely, an account is created with Channel ID and password in the ThingSpeak server which is an IoT operation platform. An application program interface (API) present in the channel acts as data read and write using the respective API key [12]. The write API key are then inserted along with other network credentials such as SSID (Service Set Identifier) network name, password and server address into the Arduino UNO IDE. The Arduino UNO in-turn unites itself with the Wi-Fi module Node-MCU and creates a connection to the said network. It then executes the source code, establishes

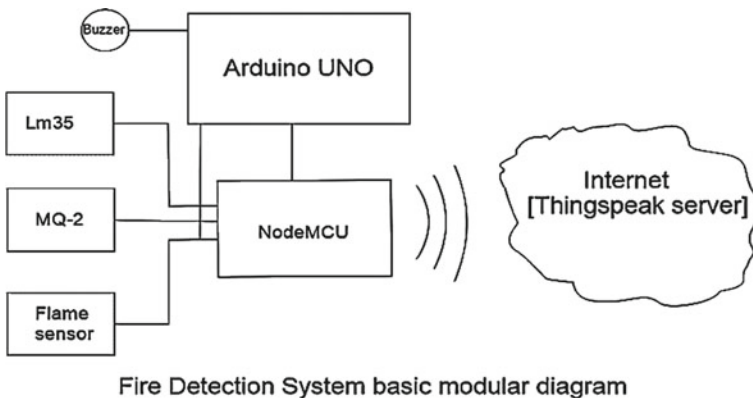
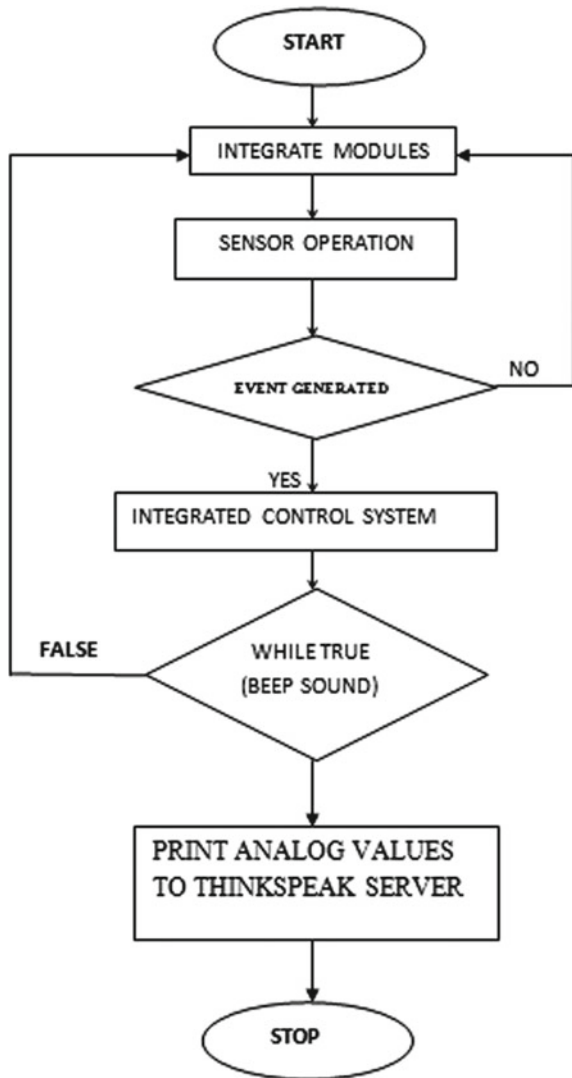


Fig. 1 Mechanism for fire detection system in simulator

communication with the Thing speak server to represent the analog data. The serial monitor present in the Arduino UNO board captures various analog data received from sensors used in the system and produces graphical format based on the input, if the said network executes properly. The graphical format is generated from the analog data, received from the sensors are later described with accurate values.

The proposed smart fire detection response system is aimed to provide the evacuation protection and dependability shown in Fig. 2 with the flowchart. It basically describes the integration of various sensors which generates a framework called

Fig. 2 Flowchart for fire detection Module



Integrated Control System. The control system when generates events will produce output values in the server.

4 Experimental Setup

4.1 Hardware and Software Components

The following are the hardware and software components used in the proposed framework.

I. NodeMCU: The Node MCU an open source platform is programmed using Arduino-Uno [13] Integrated Development Environment. It has a built in WiFi module that runs on ESP8266 WiFi and can simply connect it to WiFi with design and programming.

II. Arduino Board: The Arduino board is an 8-bit microcontroller consisting of 14 digital I/O pins that are used as either input or output in the board which are been connected with wires to various external devices and components. It consist of board voltage regulator and crystal oscillator and USB serial adapter with which the Arduino board can be design and programmed in a very effective way. The programming is prepared by means of Arduino coding or programming language (APL) that tenders 'wiring' and the Arduino Development environment (ADE) that tenders 'processing' [12]. The Arduino UNO boards are available in various sizes, system aspects, having various number of input-output pins , as for different power supply necessities, connectivity and various other applications. Arduino UNO board is a popular and reasonable of all the Arduino board based on ATmega328P microcontroller.

III. LM35 Temperature Sensor: The LM35 sensor is precision integrated-circuit devices which have an output voltage proportional to the temperature in degree Celsius [13]. It does not require any external adjustment circuitry. It is a three-terminal sensor used to measure nearby temperature ranging from -55°C to 150°C . The output voltages raises as the temperature increases.

IV. MQ Sensor -2 Smoke: The MQ-2 smoke sensor has high response time and is susceptible to smoke and to the inflammable gases such as LPG, Methane, Hydrogen etc. [13]. The sensor reads the analog signal from the Arduino and generates analog output signal. There are various types of gas sensors available based on gas type and ranges.

V. Flame sensor: The flame sensor is very sensitive and is used to detect fire source and other related fire causes. It outputs the result in the binary form 0 and 1. It is basically used for short range fire detection and is mostly accurate up to about 3 feet.

VI. Buzzer module YL-44: The YL-44 is a buzzer module which operates at 2 kHz frequency audible range. It is an active buzzer, which means that it produces sound by itself and does not need any external frequency generator. The buzzer will turn on when the I/O is kept as low and taking will turn off when the pin is kept as high. The buzzer module can be measured by Pulse width modulation.

VII ThingSpeak Graphic Interface: Thing speak graphic interface is a web-based open Application Interface (API) IoT source information platform that basically performs storing of sensed data and in turn computes those data in graphical form at the web level. It works with internet connection, as it acts as data packet carrier between the connected things. Thing speak cloud also stores, retrieves, saves, analyse, observes and work on the sensed data from the connected sensor to the host microcontroller such as Arduino. The IoT runs on comprehensive range of embedded devices and web services. The IoT can also create sensor-logging requests, location-tracking applications, and a social network of things with status updates, so that the home can be managed centrally with correct location. The main element of ThingSpeak activity is the channel, which contains data fields, location fields, and a status field. The ThingSpeak send data to the cloud where the channel stored data in the form of either private or public channel. We can write data to the channel, process and view the data with MATLAB code, and react to the data with tweets and other alerts [14, 15].

4.2 Implementing Fire Response System (Physical Structure)

The System is divided into three different modules (Fig. 3):

- 1. Temperature Module

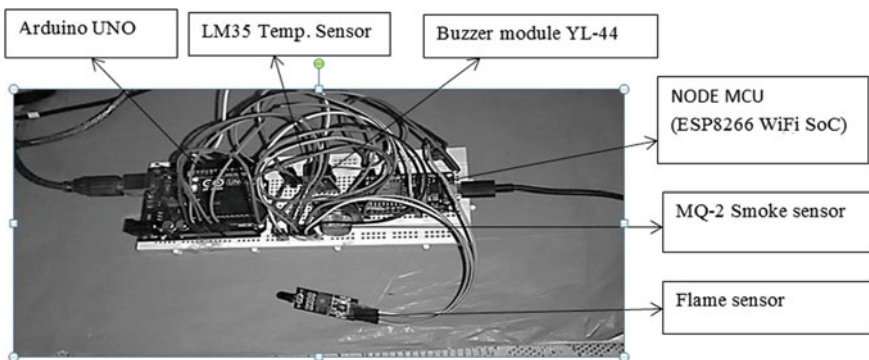


Fig. 3 Connection of Node-MCU, Arduino UNO Board, sensors with wires and cables

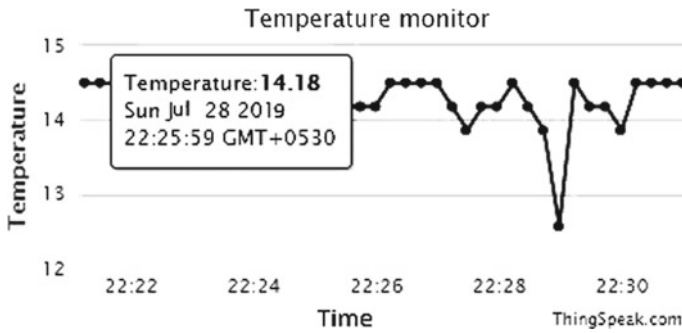


Fig. 4 Normal temperature of room with time and date

The First module is based on LM35 temperature sensor which has a detection range of -55°C to $+150^{\circ}\text{C}$. The sensor will detect rapid rise in temperature in case of fire i.e. when the temperature exceeds the fixed threshold value. Node MCU sends a message “Temperature rising” and trigger a primary alarm that will notify any person present near the target area by beeping simultaneously with a delay of 3–4 s.

Figure 4 Depicts the reading of normal temperature in a room. With the help of appropriate algorithm and respective source code the LM35 Temperature Sensor senses the normal temperature from the Arduino UNO board using wifi module Node-MCU and puts the data on the serial monitor of Arduino UNO board. The Serial monitor connects the server and place the data in the channel which reflects in the ThingSpeak server as graphical format. The temperature changes with the increase in temperature which is been implemented and verified.

1. Smoke Module

The smoke module is based on the MQ-2 Smoke sensor which detects the smoke preceding the fire; this triggers a higher magnitude of alarm with a lesser delay also relaying another message (“Smoke Detected”) via Node MCU.

In this module smoke is being detected by the MQ-2 Smoke sensor when any sort of smoke is present in the surroundings. The **MQ2** is a sensor that allows gas or smoke detector at concentrations of 300 ppm to 10,000 ppm. After calibration, the MQ-2 can detect various gases such as LPG, i-butane, propane, methane, hydrogen and fumes. It is designed for indoor use at room temperature. When the sensor detects the smoke, it becomes equals to one (smoke == 1) and the updated Twitter Status shows the message that “Smoke is Detected”.

2. Flame Module

The third and final module is based on the IR Flame sensor; it detects IR (Infra-Red) light wavelengths between 760 nm and 1100 nm (nano meter) that is emitted from fire/flames from an approximate distance of 2 m [6]. If flames are detected by the module then an even higher pitched alarm is activated with an even lesser delay of 20 ms along with the undesirable (“Fire Detected”) message.

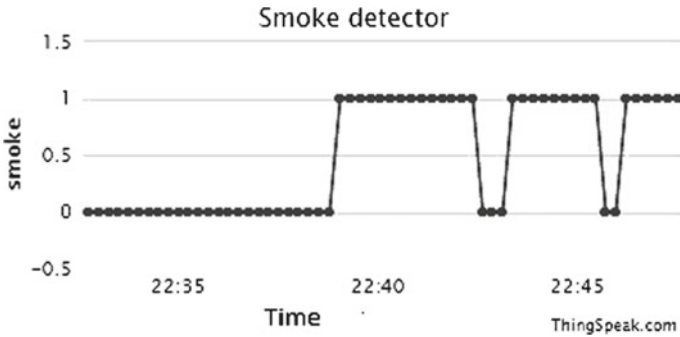
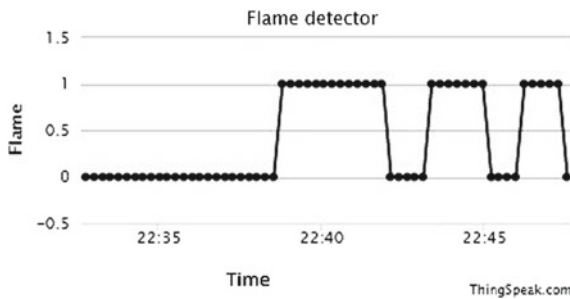


Fig. 5 Sensitivity of MQ2 sensor



In this module, the Arduino automatically **turns on** LED and Buzzer when we place a **flame** near the flame Sensor. In the Fig. 5 it is shown that the flame remains at one (Flame = 1) when it detects any kind of spark in its surroundings and comes down to Zero (Flame = 0), when the flame is removed. The Arduino UNO automatically **turns off** LED and buzzer.

5 Results and Discussion

The graphical analysis of temperature, smoke and flame detection has been described with proper readings based on the data inputs and the overall outputs is shown with different graphs as mentioned in the above figures, stating how the module accepts various inputs. The various sensors along with wifi module Node MCU and Arduino Board are embedded in a bread board and are integrated with the software that regulates the entire module and helps to process the input data supplied as temperature, smoke and flame detector which in turn responds to system. The module interacts with the data feed into the system as inputs through end-users in the Arduino Board are then send via serial monitor to the ThingSpeak Server which are then represented as outputs in different graphical forms and states the current status of the environment based on the inputs. The above module illustrates the implementation of

fire response emergency system based on the sensors range and capacity. The same module can cover large environment if the ranges of the sensors are increased to a larger strength. The module illustrates the implementation of fire response emergency system using IoT as one of the platform, which is a new area of research in terms of Security measures implemented in the field of IoT assisting various commercial and domestic purposes, meeting the requirements in today's era of science and technology through IoT enabled devices.

6 Conclusions

With the advanced and extended features of IPv6, internet of things can easily accommodate increasing number of sensors connecting to the internet. In this paper we have design a prototype for fire response system with the help of various sensors and programming features. The fire detection system use sensors like temperature, flame and smoke and these sensors behaviours to environmental inputs and are then analyzed via Thingspeak server. Several analog data are taken from the local environment in various forms and is being represented in graphical format through ThingSpeak cloud via internet. Due to scalability and addressing space in IPv6, these extended features also helps a great in carrying out the proposed fire response system in a better way keeping in mind it's security measures. The system thus tries to bring forward the application of IoT using IPv6 protocol for a fire response system. As IPv6 has a vast address space and scalability, research in IoT has a great impact in today's domestic and industrial area.

References

1. G. Ajith, J. Sudarsaun, S. Dhilipan Arvind, R. Sugumar, IoT based fire deduction and safety navigation system. *Int. J. Innov. Res. Sci. Eng. Technol.* (2018)
2. M. Chernyshev, Z. Baig, O. Bello, S. Zeadally, Internet of Things (IoT): research, simulators, and testbeds. *IEEE Internet Things J.* (University of Kentucky, USA, vol. 5, issue 3, June 2018)
3. S. Ziegler, C. Crettaz, L. Ladid, S. Krco, B. Pokric, A.F. Skarmeta, A. Jara, W. Kastner, M. Jung, *IoT6 – Moving to an IPv6-Based Future IoT* (2016)
4. H. Safa, S. Priyanka, V.G. Priya, B. Vishnupriya, IOT based Theft preemption and security system. *Int. J. Innov. Res. Sci. Eng. Technol.* (2016)
5. Q. Ding, Z. Peng, T. Liu, Q. Tong, *Multi-Sensor Building Fire Alarm System with Information Fusion Technology Based on D-S Evidence Theory* (2014)
6. C.-S. Ryu, *IoT-based Intelligent for Fire Emergency Response Systems* "Department of Cartoon & Game Animation, Yewon Arts University, Seoul, pp. 482–863, Korea, International (2015)
7. S. Singh, S. Sharma, D. Garg, Ganeshan, *Intelligent Fire Emergency Response System*, Department of Computer Science Engineering, SRM University, Chennai (2017)
8. C. Malarvizhi, V. Kalaipoonguzhali, J. Anitha, *Microcontroller ATmega 328P and GSM Based Advanced Home Security System* (2017)
9. R. Angeline, S. Adithya, A. Narayanan, Fire alarm system using IOT. *Int. J. Innov. Technol. Exploring Eng.* (IJITEE) ISSN: 2278-3075, vol. 8, issue 6S3, April 2019

10. M.S.B. Bahrudin, R.A.K.N. Buniyamin, Development of fire alarm system using Raspberry Pi and Arduino Uno, in *2013 International Conference on Electrical, Electronics and System Engineering*
11. R. Chutia, D. Sonowal, S. Sharma, Remote household appliance control system using GSM, in *Proc. of the International Conference on Advanced Computing and Communication Technologies (ACCT 2011)* ISBN: 978-981-08-7932-7
12. G.M.S.M. Rana, A.A.M. Khan, M.N. Hoque, A.F. Mitul, Design and implementation of a GSM based remote home security and appliance control system, in *Proceedings of the 2nd International Conference on Advances in Electrical Engineering*, Dhaka, 19–21 December 2013, pp. 291–295
13. M.S.B. Bahrudin, R.A.S.N. Buniyamin, *Development of Fire Alarm System using Raspberry Pi and Arduino Uno* (2013)
14. N. Sinha, K.E. Pujitha, J.S.R. Alex, Xively based sensing and monitoring system for IoT, in *International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, pp. 1–6 (2015)
15. S. Pasha, Thingspeak based sensing and monitoring system for IoT with Matlab analysis. *Int. J. New Technol. Res. (IJNTR)* ISSN: 2454-4116, vol. 2, issue 6, June 2016

Intelligent Physical Access Control System Through Three-Stage Verification Using IoT



Rohit Tripathi, Manoj K. Shukla, and Yogesh Kumar

Abstract In present communication, an intelligent physical access control system is using IoT platform, where there are three different verification stages. The hardware is designed on proteus and then fabricated on PCB board. The major part of this system is Arduino Uno, fingerprint scanner-R305, 4×4 keypad module, GSM module-900 A, 5 pin relay-5 V, and solenoid lock. Arduino IDE software is used to operate the hardware. Five steps have been accounted, first: enable the system, second: put finger on scanner, if it matches, third: asking to enter password or code for second verification then fourth: one enters the password and access or lock is opened. In meanwhile, one SMS delivers on owner/authorized person. Fifth: after certain delay (5 s), it locked, automatically and again one more SMS delivers on owner/authorized person. Such systems are designed and applied for higher security purpose for highly restricted areas for common public or hotels, classrooms of school and colleges, offices, bank lockers where unauthorized person is not permitted to access the place. The total cost of this hardware is approximately Rs. 2700.

Keywords Smart lock · Arduino · IoT · Three protocols

1 Introduction

Security systems are essentials, which provide privacy at their own point of requirement. It can be simple like a secured access system or a complex one like: smart home automation system, which takes a large no. of different security parameters over continuous changing technology. Security is the biggest issue in any individual and industrial whether it is in the personal or social way. Security of human life and property is one of the paramount challenges facing any nation or any corporate organization. Security systems are necessary everywhere especially in banks,

R. Tripathi (✉) · Y. Kumar
Electronics and Communication Engineering, Galgotias University, Noida, India

M. K. Shukla
Electronics Engineering, Rajkiya Engineering College, Kannauj, UP, India

houses, offices etc. Now, security systems are essentials which lead optimum use of the resources and safety in society. A security system must lead to different step of authentication, only a door lock cannot provide us security at its most, so some other features like biometric-based and many more approaches are highly required for smart security systems. As the technology is evolving every second, intelligent security systems have been developed and implemented to prevent illegal invasion.

The aim of this study is to present most secure access system with having multiple verification stages for genuine person access only. Before gaining access, it is categorized into single and multiple protocol system depending on the security huddle. First step toward security was lock and key system. And security protocol followed in this system was “single key for a single lock”. But after that many new featured security or access control systems have been introduced as per literature.

The circuit has been designed for intelligent door lock where this equipment was designed as an intelligent door which locked the access easier than earlier existing models. It could set user cards for the door lock, synchronize the door lock time, read the door-open record, and search the door-open record according to the door-open time or the card number [1]. GUI-based face recognition system for door lock control using microcontroller was developed [2]. It showed that system was developed to become fully automatic face recognition system by adding another feature that automatically captures image with a sensor when there is human presence at the entrance area. The new design through smartphone NFC door lock control system was developed which provided a low power, secure, convenient way both to open and to lock a door [3]. One access control system was developed based on e-mail alert along with a sophisticated design which provided the flexibility of remote access control while ensuring security. In one other research, RFID-based automatic door locking and unlocking system has been designed where Arduino was used as microcontroller which given the user data in real-time system [4, 5]. Motion and image capturing-based smart access control on Android phone was also developed [6]. Utilizing of both input parameters achieved another advantages to build the proficiency regarding unique identification in premises.

Smart door lock system based on blockchain was also developed where the blockchain was applied to the smart access control system which also guaranteed authentication, non-repudiation, and data integrity. It prevented an unauthenticated user from participating in the blockchain network [7]. A design of lock system was developed for operating door, and it used Bluetooth technology with low power and is available on almost all gadgets [8, 9]. Raspberry Pi 3-based vehicle door locking system was powerful system with high security. This designed system could not be cracked by unauthorized persons. Car locking was only done by genuine finger identification only [10]. A safe door lock system based on Raspberry Pi was developed where cameras, keypad, and Pi-lids were being utilized to provide an alarming system that had the ability to notify the owner, as well as, recognizing guests by giving them a user id [11, 12].

One other system was also designed as door lock system which was combined with augmented reality technology (AR) and mobile phone remote control system. But the demerits of using AR technology were that the speed of opening the door was

slower than using the real lock key [13]. Smart lock system using crypto-algorithms was also presented where a complete system was developed with an Android app, using cryptographic algorithms for secure transmission and hardware with actuators in order to control unauthorized access [14, 15]. One latest smart lock system was presented where a voice of human was detected by FET and then this voices processed by combining MATLAB and Arduino. Through this, this voice was being processed and a door lock system was operated through Arduino [16].

Password-based security system and fingerprint-based security system are some of the examples of single protocol system. A multiprotocol system is also made with incorporation of more than one level of protocol system. The protocol system solely represents the advanced locking system, and it can be two factors-based or may be three factor-based security system, depending upon the requirement of the individual.

At present study, the designed system provides multilevels authentication in the real time, and with the real time which makes the highest secure system. A combination of fingerprint and password-based locking features on a single GSM module produces opening and closing of the lock as output on the valid input, and the GSM module provides messages to owner/authorized person.

2 Internet of Things (IoT)

Internet of things (IoT) is an advanced automation and analytics system which exploits networking, sensing, big data, and artificial intelligence technology to deliver complete systems for a product or service. These systems allow greater transparency, control, and performance when applied to any system. The most important features of IoT include connectivity, sensors, active engagement, and small device use. The main components of IoT-based system are microcontroller like: Arduino or Resberry Pi, sensors and actuators which help the designed system to operate wirelessly with the availability of Internet only. The application of IoT exits across all smart lifestyle and businesses or industries. IoT involvement offers improved customer engagement, technology optimization, reduced waste, enhanced data collection, remote operation, and real-time monitoring. In this system, Arduino has been taken as microcontroller due to wide availability and low cost.

3 System Description and Principal

In this study, a multipurpose and multilevel verification-based access control system has been designed. Before designing the present system, the detailed review has also been carried out. It is found that Arduino is best microcontroller for IoT application and it is widely available and having low cost. After that, Arduino has been taken. Here, three-level verification protocol has been adopted, where first verification stage is fingerprint through fingerprint scanner, second one is password through keypad

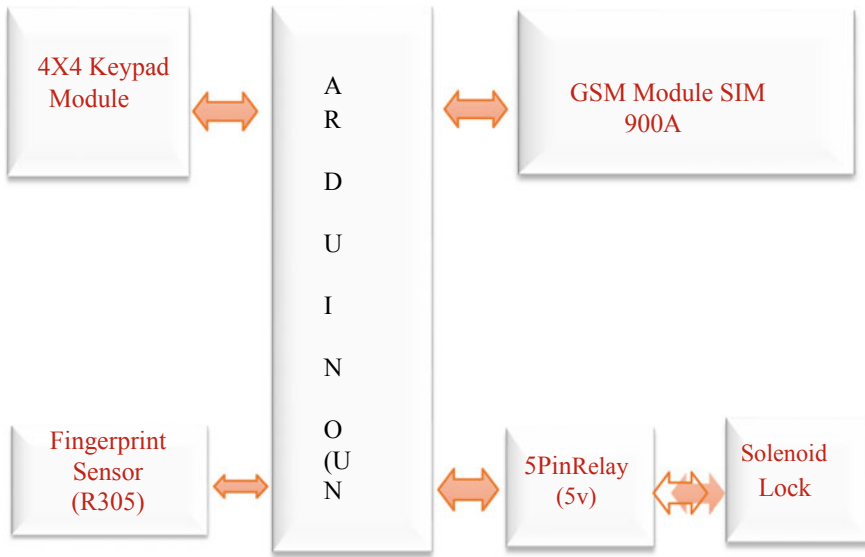


Fig. 1 Block diagram of proposed multipurpose door locking system through Arduino

and if access is allowed after both verification stages, sms has been delivered to owner or authorized person as alert in both access on as well as access off stage. The block diagram of the system has been shown in Fig. 1.

The locking system setup mainly consists of Arduino UNO. It also contains of 4×4 keypad, fingerprint sensor (R305), GSM module (900 A), solenoid lock, jumper wires, etc. The Arduino IDE has been used as software and uses code in C++ language. Arduino Uno H/W board has been connected to PC through connecting cable. The circuit diagram of the designed system has been shown in Fig. 2.

Once the hardware and software is connected, the “enroll” code is uploaded and then the fingerprints for the input purpose are taken from human finger with commands on the serial monitor and stored for the further verification purpose.

After the installation of all the components, the enroll code is entered and uploaded to the Arduino Uno board and then the fingerprints have been taken and enrolled from multiple users. Then after the final code is uploaded to the board as a password/key code to be entered and function code for the process, the hardware is ready for working in proper manner with accessible only for authorized persons who can get access in particular location or area. Such systems can be implemented, where restricted areas are unsafe for general public like: bank lockers, confidential cupboards or rooms, VIP areas, or any highly restricted areas.

Now, process starts as any human put his/her selected finger on fingerprint scanner, if it is correct then green LED glows and next verification enables after that enters the predecided code or pass word through keypad, if it matched the code, access is ON state and at the same time, sms is delivered to owner/authorized person’s mobile as alert only. After 5 s, automatically this access has been OFF and sms is delivered to

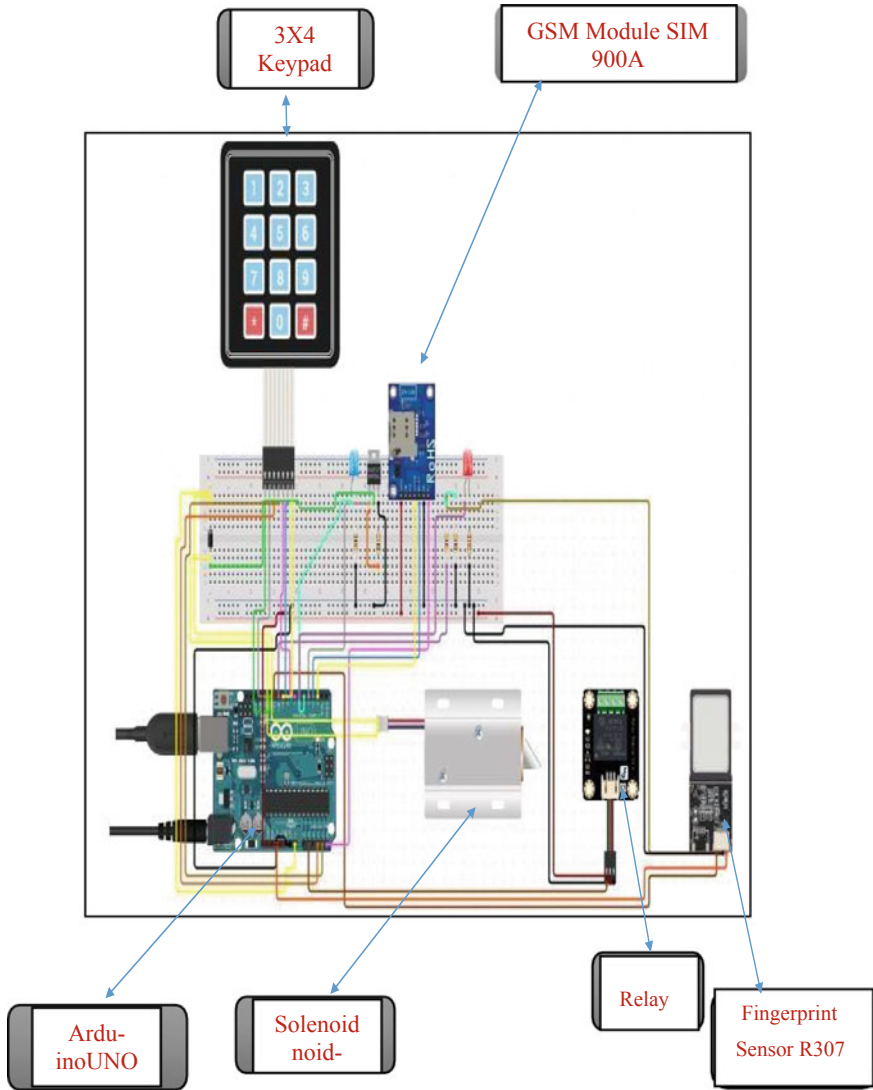
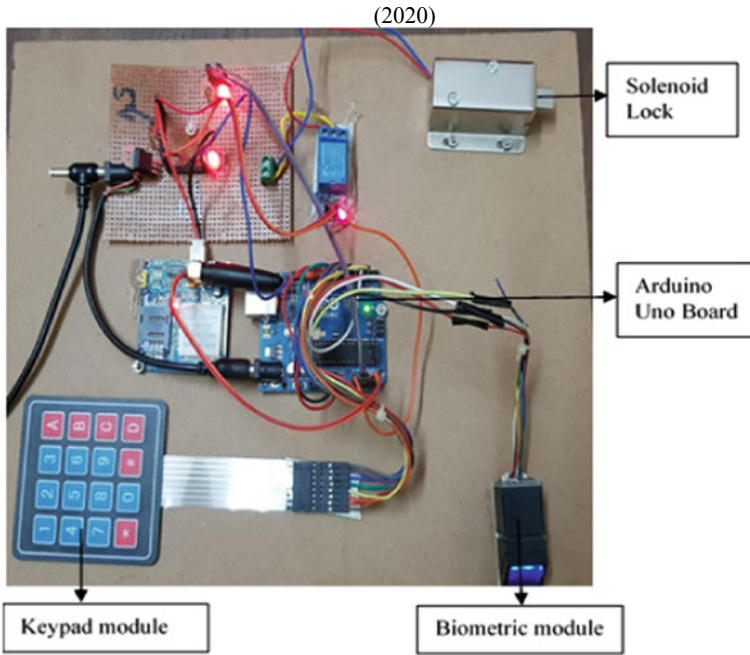


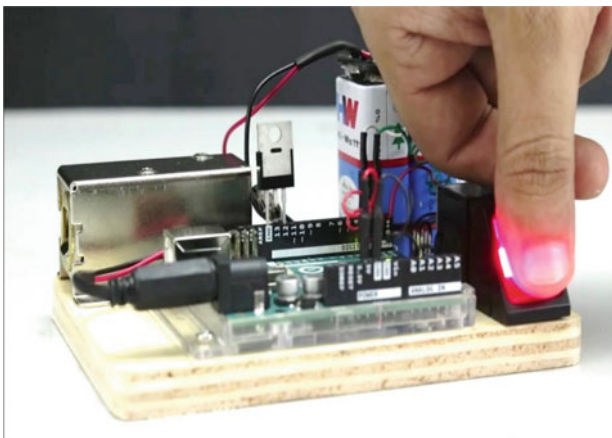
Fig. 2 Circuit diagram of proposed multipurpose door locking system through Arduino

the same person as alert. At any stage, fingerprint or code does not meet the precodes, the system does not allow to go forward and lock will not be opened.

Following above methodology, it is called three-stage verification access control system which gives higher level of security and alert system. The flowchart of the process and the picture of hardware circuit have also been shown in Figs. 3 and 4.



(a)



(b)

Fig. 3 a, b Actual hardware picture of proposed multipurpose door locking system through Arduino Uno

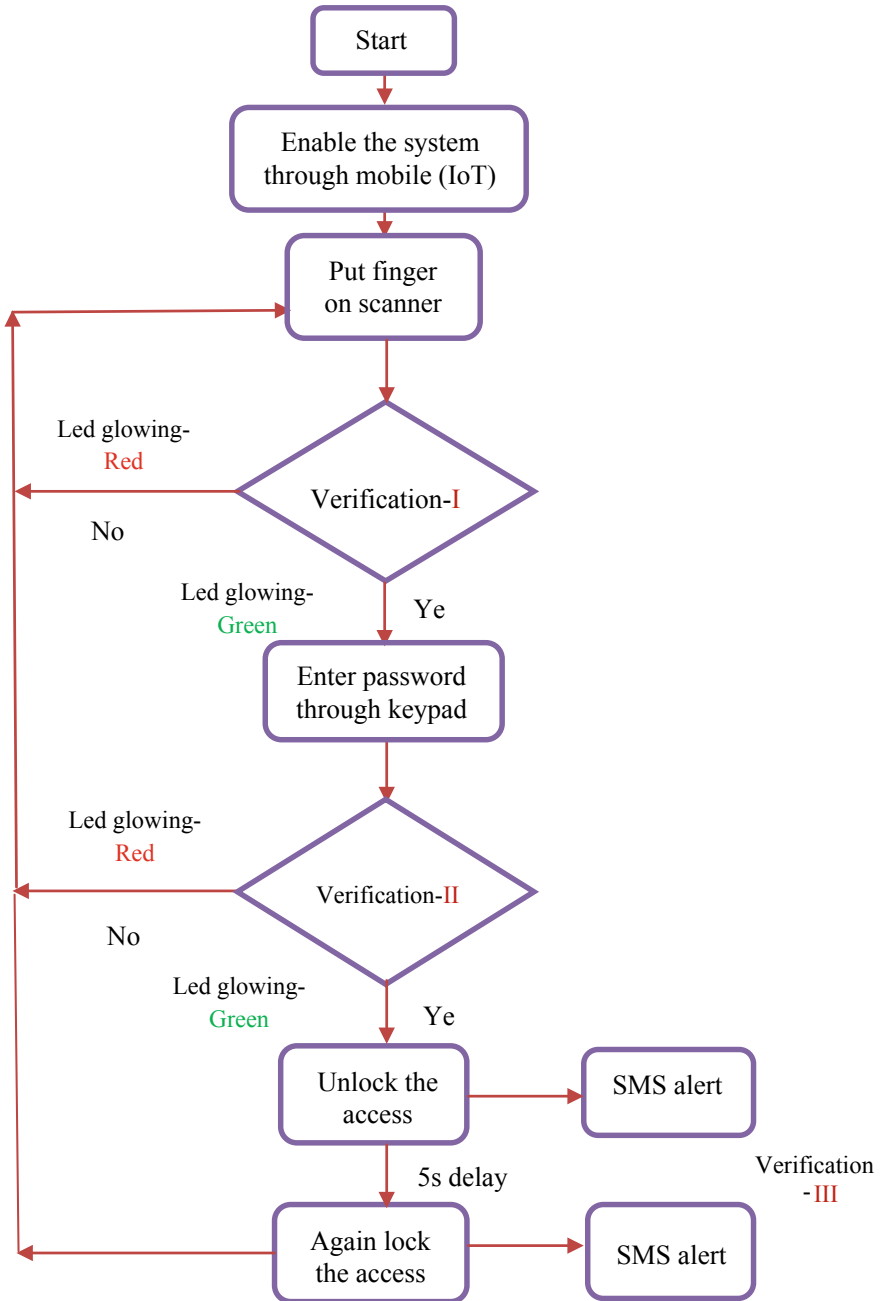


Fig. 4 Flowchart of system working through two-stage verifications

4 Methodology

See Fig. 4.

5 Results and Discussion

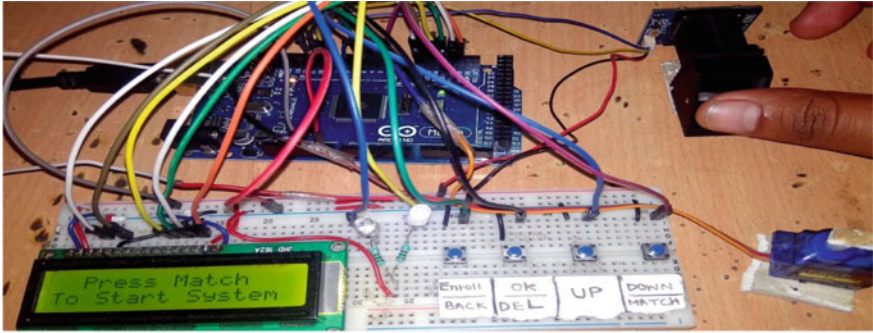
After fabricating the system, there are five steps in the process. First: enable the system, second: put finger on scanner, if it matches, third: asking to enter password or code for second verification then fourth: one enters the password and access or lock is opened. In meanwhile, one sms delivers on owner/authorized person as “YOUR LOCK IS OPENED ON DD-MM-YYYY AT HH:MM”. Fifth: after certain delay (5 s.), it again locked, automatically and again one more sms delivers on owner/authorized person as “YOUR LOCK IS CLOSED ON DD-MM-YYYY AT HH:MM” and process completes. The pictures of all five steps of hardware have been shown in Figs. 5a–e.

These three levels provide us with the security feature we can demand for prevailing in the current, security conditions around us. It is easy to design and easily accessible for all type of users.

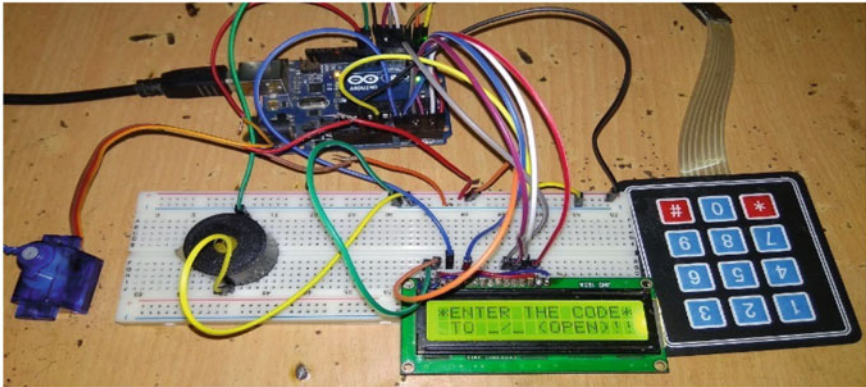
6 Conclusions

An intelligent locking system is presented here, which gives us clear-cut information about the different approaches which is required to take in consideration about the different aspects of security systems. With the increasing threat and burglary through advanced techniques, now is the time to adapt for modern solutions, the modern solutions list different approaches of protocol system like password-based, face detection, and biometric-based security systems.

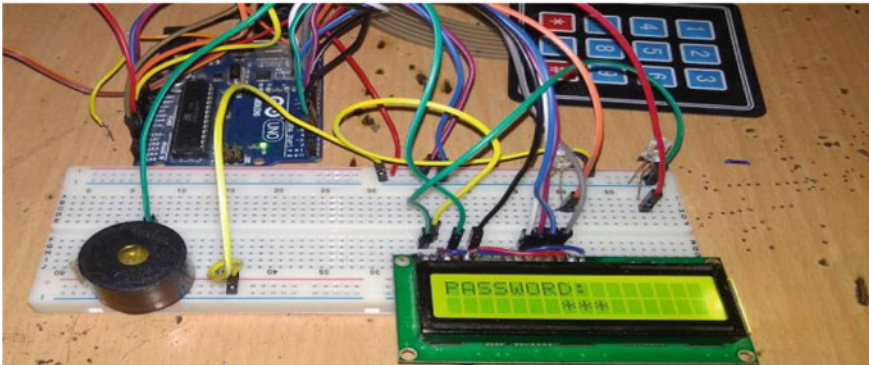
The approach is combining with password and fingerprint-based security system, with a GSM module in combination, which gives us messages on our phone when the system is being locked or unlocked. It can be easily applicable to get access of our home lock security system, being on a far remote place too. The message feature provides an alert on instant basis and we can stop the authorized person through mobile app with NO command or DND command through mobile also.



(a)

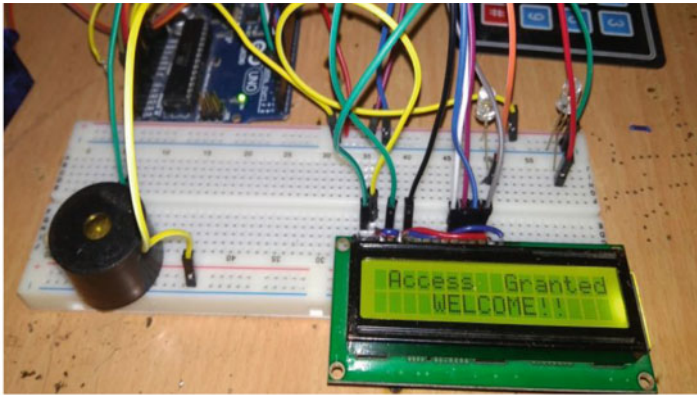


(b)

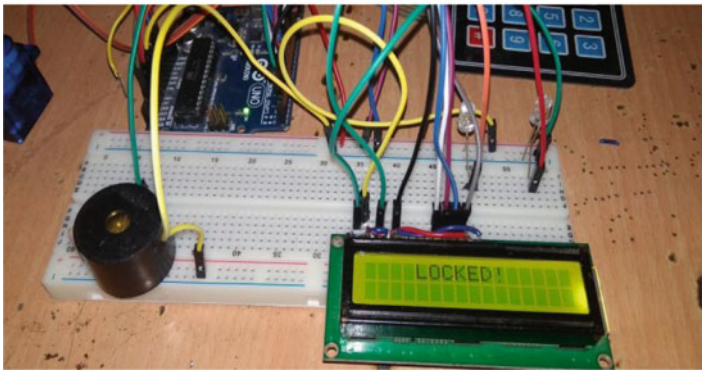


(c)

Fig. 5 a–e Complete hardware connected with PC



(d)



(e)

Fig. 5 (continued)

References

1. A. Fang, X. Ye, M. Gong, W. Yang, L. Zhang, The circuit designing of the intelligent door lock, in *IEEE International Colloquium on Computing, Communication, Control and Management, Guangzhou, China*, pp. 778–781 (2008)
2. H. Hassan, R.A. Bakar, A.T.F. Mokhtar, Face recognition based on auto-switching magnetic door lock system using microcontroller, in *IEEE International Conference on System Engineering and Technology, Bandung, Indonesia*, pp. 1–6 (2012)
3. C.H. Hung, Y.W. Bai, J.H. Ren, Design and implementation of a single button operation for a door lock control system based on a near field communication of a Smartphone, in *IEEE 5th International Conference on Consumer Electronics Berlin (ICCE-Berlin)*, pp. 260–261 (2015)
4. R.K. Gupta, S. Balamurugan, K. Aroul, R. Marimuthu, IoT based door entry system. *Indian J. Sci. Technol.* **9**(37):1–5
5. S. Nath, P. Banerjee, R.N. Biswas, S.K. Mitra, M.K. Naskar, Arduino based door unlocking system with real time control, in *2nd IEEE International Conference on Contemporary Computing and Informatics (IC3I)*, pp. 358–362 (2016)

6. F.C.A. Aman, Motion sensing and image capturing based smart door system on android platform, in *IEEE International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, Chennai, pp. 2346–2350 (2017)
7. D. Han, H. Kim, J. Jang, Blockchain based smart door lock system, in *IEEE International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, pp. 1165–1167 (2017)
8. M.S. Hadis, E. Palantei, A.A. Ilham, A. Hendra, Design of smart lock system for doors with special features using bluetooth technology, in *IEEE International Conference on Information and Communications Technology (ICOIACT)*, Yogyakarta, pp. 396–400 (2018)
9. O. Agbo David, M. Chinaza, O. Jotham, Design and implementation of a door locking system using android app. *Int. J. Sci. Technol. Res.* **6**(8), 198–203 (2017)
10. N.N. Nagamma, M.V. Lakshmaiah, T. Narmada, Raspberry Pi based biometric authentication vehicle door locking system, in *IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, Chennai, pp. 2348–2351 (2017)
11. N.A. Hussein, I.A. Mansoor, Smart door system for home security using Raspberry pi3, in *International Conference on Computer and Applications (ICCA)*, Doha, pp. 395–399 (2017)
12. R. Kaushik, T. Arora, T.R. Sukanya, Design of Eyewriter for ALS patients through Eyecan, in *IEEE International Conference on Advances in Computing, Communication Control and Networking, India*, pp. 991–995 (2018)
13. R. Tripathi, Nitin, H. Pratap, M. K. Shukla (2021). Intelligent car sabin safety system through IoT application. In: Sharma, H., Saraswat, M., Yadav, A., Kim, J.H., Bansal, J.C. (Eds.) *Congress on Intelligent Systems, Proceedings of CIS 2020, Vol 1. Advances in Intelligent Systems and Computing*, Springer, Singapore.
14. T. Gupta, R. Tripathi, M.K. Shukla, S. Mishra, Design and development of IOT based smart library using line follower robot. *Int. J. Emerging Technol.* **11**(2), 1105–1109 (2020)
15. M. Ahtsham, H.Y. Yan, U. Ali, IoT based door lock surveillance system using cryptographic algorithms, in *IEEE International Conference on Networking, Sensing and Control (ICNSC)*, Banff, AB, Canada, pp. 448–453 (2019)
16. L.D.W. Raj, K. Santhosh, S. Subash, C. Sujin, P. Tharun, Voice controlled door lock system using Matlab and Arduino, in *IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*, Pondicherry, India, pp. 1–6 (2019)

Descriptive Indic Answer Script Evaluation Using Deep Learning



Deepak Senapati , Abinash Panda , Siddhant Mund ,
and Sunil Kumar Mohapatra 

Abstract The evaluation of descriptive answer scripts is a monotonous, time-intensive process and is prone to human errors and sometimes bias. As a result, a student's career can be at stake. Hence, in order to mitigate the errors and efforts during answer script evaluation, an automated evaluation system is proposed to assign marks to the answer scripts correctly. Our proposed system, 'Descriptive Indic Answer Script Evaluation using Deep Learning' will provide the marks to an answer based on its content and correctness depending upon the model trained using natural language processing and deep learning techniques on the dataset containing some of the manually evaluated sample answers. In the proposed sequential model, the learning is done by the convolutional and LSTM cells. The input to the system is the extracted text from the answer script which after preprocessing was fed to the model for training. As there was no significant system available for Indic languages, the system will be very helpful for the education boards conducting examinations in Indic languages. In this paper, among the different Indic languages, we have focussed on Odia language as very minuscule amount of work has been done in this language with regard to any NLP task.

Keywords Convolution neural networks (CNN) · Deep learning · Indic languages · Long short-term memory (LSTM) · Natural language processing (NLP)

1 Introduction

In every educational institution, examination is done to assess what students have learned with regard to particular subjects. The evaluation of answer sheets is a very important task because a student's ability is judged by the marks obtained by him/her in the exam. Manual answer sheet evaluation requires a lot of time and concentration. Sometimes, there can be mistakes as the evaluator has to evaluate heaps of answer sheets. Therefore, we proposed 'Descriptive Indic Answer Script Evaluation using

D. Senapati · A. Panda (✉) · S. Mund · S. K. Mohapatra
College of Engineering and Technology, Bhubaneswar, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_69

901

Deep Learning,' to automatically evaluate answer sheets in very less time based on the model trained upon some of the manually evaluated sample answers. We mainly focussed on Indic languages as almost negligible work has been done in this field. Automatic answer sheet evaluation in Indic language is a very new task that has been discussed in this paper. The model will evaluate the answer automatically; only, one has to provide a scanned image of the answer then the system will extract the text from the image using Optical Character Recognition (OCR), and then, after text preprocessing, the model will evaluate the answer and provide the score. As we have mainly focussed on Indic Language, this model can be implemented in various state government schools and colleges for answer sheet evaluation. The model that has been implemented is a combination of Natural Language Processing (NLP) and deep learning. Deep neural networks are able to capture semantics of text in order to detect the similarity between texts. Two models were implemented using convolutional neural network (CNN) and long short-term memory (LSTM), and based upon their performance, LSTM model was finalized.

The trained model predicts the mark of the input answer with a decent accuracy. The proposed model helps in easy evaluation of bulk answer sheets in less time. The proposed algorithm can be implemented by state government-run schools which have Indic language in their curriculum.

2 Literature Review

The field of automatic answer sheet evaluation evolved in the early 1990s. Recently, it involved natural language processing in its domain for achieving the purpose. Ellis Batten Page developed an automated evaluation system (AES) as Project Essay Grader based upon a request from the American College Board which is basically used for essay analysis. Later on, many more systems were developed which are currently in use at different educational institutions. Landauer [1], in 2003, developed a system named Intelligent Essay Evaluator where latent semantic analysis (LSA) was used for evaluation. Then Kakkonen [2] used the probabilistic LSA in his automatic answer evaluator tool in 2005. Along with LSA, other classification methods like naive Bayes theorem, maximum entropy, K-nearest neighbour had also been experimented in different works. A competition was hosted by Hewlett foundation on the Kaggle platform as the Automated Student Assessment Prize in 2012 for automated essay scoring. Naive Bayes theorem was used in 2018 by Piyush Patil [3] in his proposed system—subjective answer sheet evaluation. According to this system, the answer is evaluated in three phases as keyword matching followed by a grammatical check and then comparing with the question-specific things. Here, Bayes classifier classifies the answers among ten classes(0–9). For grading, it considered the output class along with the probability of its occurrence from Bayes classification. For example, in a marking scheme of 10, the grade will be—(obtained class + probability of class occurred/100). Similarly, Prayag Singh [4] used Boruta algorithm, random forest along with gradient boosting in his proposed model. The system

Table 1 Attributes in training dataset

Name	Type	Description
img_id	string	Unique identification for each individual answer
ans	string	The actual answer of the student
marks	float	The corresponding score

indicates usage of Boruta algorithm for an upside-down search of features for feature selection under categories of ‘Important,’ ‘Confirmed,’ ‘Rejected,’ ‘Tentative’ and the training of these subsets using the random forest for increased performance and decreased training time. Then, gradient boosting is used to develop the model from the decided features. Later in 2019, Neethu George [5] insisted on using deep learning for a better and effective result in regard to automated answer sheet evaluation. The system works with an recurrent neural network (LSTM RNN), where inputs will be temporal data from embedding layers and embedding vectors correspond to the final glove vector where the glove vector is the semantic representation of the whole answer with the use of the respective embedding layer. According to the system, the dropout layer will access the input and then will relate it to a fully connected neural network where the activation function is Softmax. The system was named as deep descriptive answer scoring model (D-DAS).

3 Dataset

Dataset plays a crucial role while assigning a score to an answer using the developed model evolved from the relation between the answers and their corresponding scores. So, we collected 100 images of a pre-evaluated handwritten answer. On performing some basic image processing operations, we got 84 effective images for consideration. We used optical character recognition (OCR) for the extraction of answers from the images. Then, the answers and their respective marks were accumulated into a CSV-formatted file structure for designing the training dataset. The CSV-formatted attributes were (Table 1):

A sample of the dataset is given in Fig. 1.

4 Preprocessing

Text preprocessing [6] is the first step in any natural language processing tasks. Strings cannot be fed into machine learning algorithms for learning purposes. Hence, the answers in our dataset need to be represented in the form of numerical features after being cleaned for any irregularities. For the text preprocessing task in the project, a library Indic natural language processing toolkit (iNLTK) was used. It is modelled

Fig. 1 Dataset

	img_id	ans	marks
0	IMG01	ଇଂରରେଜକଂ ସାମରାଜ୍ୟ କୁ ସୁଦୃଢ଼ କରିବା ପାଇଁ ବେଲସେଲି ...	5.0
1	IMG02	ଇଂରରେଜକଂ ପ୍ରଭାବିତ ରାଜ୍ୟ ବିସ୍ତାର ପାଇଁ ରାଜ୍ୟାଭିନ...	5.0
2	IMG03		NaN 3.0
3	IMG04	ସାମନ୍ତ ସହି ପ୍ରଥା ବେଲସେଲି ପ୍ରଣୟନ କରିଥିଲେ।ଏହାର...	4.0
4	IMG05	ସାମନ୍ତ ସହି ଲର୍ଡ଼ ମାକ୍‌ସଲେକଂ ଦ୍ୱାରା ପ୍ରଣୀତ ହୋଇଥି...	1.0
...
79	IMG80	ସାମନ୍ତ ସହି ବେଲସେଲି ପ୍ରଣୟନ କରିଥିଲେ ଇଂଗ୍ରେଡ଼ ମା...	4.5
80	IMG81	ଭାରତର ରାଜ୍ୟଗୁଡ଼ିକୁ ଅଧିକାର କରିବା ପାଇଁ ବେଲସେଲି ସା...	3.0
81	IMG82	ଇଂଗ୍ରେଡ଼ ମାନେ ଭାରତରେ ବାଣିଜ୍ୟ କରିବାକୁ ଆସିଥିଲେ।ଭ...	2.0
82	IMG83	ସାମନ୍ତ ସହି ବେଲସେଲି ପ୍ରଣୟନ କରିଥିଲେ ଇଂଗ୍ରେଡ଼ ମ...	4.5
83	IMG84	ଇଂଗ୍ରେଡ଼ ମାନକ ସାମ୍ରାଜ୍ୟ ବିସ୍ତାର ପାଇଁ ବେଲସେଲି...	5.0

84 rows × 3 columns

on the lines of NLTK library, but on a much smaller scale, and has support for 12 Indic languages in addition to English. Herewith are the steps undertaken for text preprocessing—removal of NaN (not a number) values from ‘ans’ column, the number of words feature extraction, vectorization, padding and trimming of the vectors.

As stated, the first step in preprocessing is the NaN (not a number) value removal. The NaN values cannot be fed into the deep learning models and also do not have any significance; thus, they were removed. On checking the ‘ans’ column, it was found that there were some NaN entries. After removing the NaN entries, it resulted in a reorganisation of the dataset index; hence, in order to get an easy working environment, the dataset index was reset. After the NaN values were removed from the ‘ans’ column, the total number of entries was reduced to 67.

Next, we move on to feature extraction, a process of creating a new feature from existing attributes. We extract the ‘number_of_words’ feature from the dataset by finding the length of the tokenize list which is obtained by splitting the sentences about a space. It was found that the number of words do not have a direct relationship with the marks secured, but it can be used to standardize the vector length after the vectorization process before being fed into the training cycles.

Coming to vectorization, it is the most important step in preprocessing. Vectorization is transforming or converting text into corresponding meaningful vectors of real numbers that a machine can understand. For carrying out the vectorization process, Indic natural language toolkit (iNLTK) library was used. iNLTK library uses transfer learning technique [7] and universal language model fine-tuning (ULMFiT) algorithm [8] for vectorization. Transfer learning is the process of using any pre-trained model and tweaking them for the problem at hand. ULMFiT is a prominent transfer learning technique implemented in the natural language processing space. ULMFiT

is trained on the Wikitext-103 corpus and uses an AWD-LSTM (ASGD weight-dropped LSTM) for its language modelling tasks. Using the ULMFiT pre-trained model and transfer learning, the iNLTK library retrained it on the Odia Wikipedia corpus for the embedding vectors. The iNLTK vectorization method generates an output of 1×400 vector for each token. Vectorization for the entire dataset generates a 3D vector of shape (67, no_of_tokens, 400). The first dimension is the number of samples in the data frame, second dimension is the number of tokens in each sample, and the third dimension is the vector length. Thus, for each sample, a 2D vector is generated given by (no_of_tokens, 400).

Given that the no_of_tokens is variable from sample to sample, thus, it cannot be directly fed into the neural network for training. This brings us to the final step in preprocessing—padding and trimming of the vectors. For padding, the vectors were post-appended with 1×400 zero vectors, and for trimming, the vectors were trimmed from the left end. To decide upon the final length of the vectors, the feature ‘number_of_words’ was used which was extracted earlier. To decide upon the vector length, the statistical distribution of the number_of_words feature was computed. It was observed that the max length was 211 and the median length was 124 with the 75 percentile mark at 156 words. For padding, the median length of the vectors was considered to avoid outliers. Padding was done using the pad_sequence() from Keras library. This function pads a vector with the given series to get the specified length and trims the vectors in case it is longer than the specified length. This resulted in a NumPy array of shape (67, 124, 400) which would be fed into the deep learning models for training.

5 Deep Learning Models

Two deep learning models were proposed—CNN and LSTM—with a convolution layer. Training was done for both the networks, and based upon performance, the best among them was selected.

CNN Model

Convolutional neural network or ConvNets (CNN) is a popular deep learning model used to perform classification tasks from images, text, videos or sound [9, 10]. We proposed a CNN model with the vectorized Odia text tensor as input and multiple hidden layers. The output of the model was an array of length 6.

The layers used in the sequential model are:

- Input layer—1D convolution layer
- Hidden layer—MaxPooling1D layer, flattening layer, dense layer
- Output layer—fully connected dense layer.

Conv1D. The Conv1D creates a convolutional layer. A dropout of 0.2 was added to the sequential model which helps in preventing overfitting of the data by randomly

switching off some of the nodes. Sixty-four nodes were created in the convolution layer of the sequential model. `Input_shape` of [124, 400] was used; as mentioned earlier, the padded length of the vectors is 124 where each token's vector length is 400. We use ReLu nonlinearity as activation, and stride was set to 1.

MaxPooling1D. The MaxPooling1D creates a max pooling layer, the only argument is the pool size [11]. The pool size of 2 was used as it is the most common. As it is known that the stride length and window size are the same by default, thus it was 2.

Flatten. After the convolution + pooling layers, we flatten their output to feed into the fully connected layers.

Dense Layers. The dense function in Keras constructs a fully connected neural network layer, automatically initializing the weights as biases [12]. The number of nodes in the hidden dense layer is 60, and the activation function that we used for non-linearity is ReLu. The unit size of the output dense layer is 6, and as discussed in data preprocessing, there are six unique values of marks in the dataset.

The sequential model was compiled using the Keras API compile function with Adam optimizer as the optimizer function, sparse categorical cross-entropy for loss computation and accuracy as the evaluation metric of the model (Fig. 2).

LSTM Model

LSTMs or Long short-term memory networks are a special type of recurrent neural networks (RNNs) [13] and are capable of learning long-term dependencies. We proposed an LSTM model [14] with a convolution layer in the beginning which would help in reducing the dimensions of our input data and thus also reduce the cost of training.

A sequential model was proposed with the following layers:

- Input layer—1D convolution layer
- Hidden layer—one LSTM layer and one fully connected dense layer
- Output layer—fully connected dense layer

1-D Convolutional layer. The filters are the number of simultaneous filter convolutions carried out. It specifies the dimensionality of the output space. The number of filters used were equal to `pad_length` which is 124. The output of this layer is a 2D vector of shape— (85, 124).

LSTM Layer. LSTM layer is the most important layer of the model. This layer consists of 96 nodes, and a dropout of 0.2 was added. In this case, 20% of the total nodes were switched off in each training cycle. The output of the LSTM is a 1D vector with shape (96).

Hidden Dense layer. A fully connected dense layer was used to accept the output of the LSTM layer. This layer has 24 nodes, and a 'ReLu' activation function was applied. ReLU stands for rectified linear unit, and mathematically, it is defined as $y = \max(0, x)$. This is used as it is proven to be a really fast function as it converges

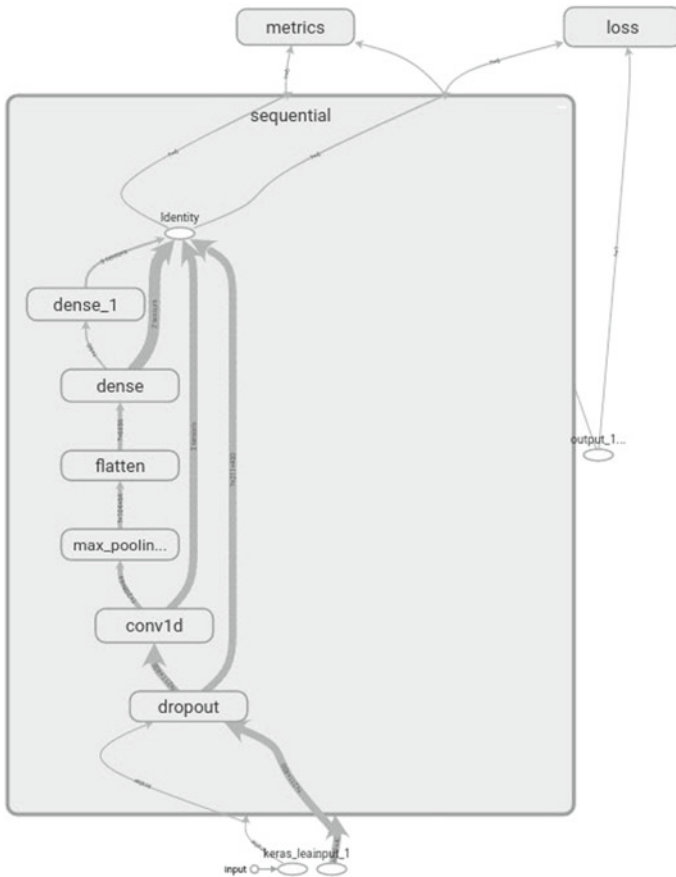


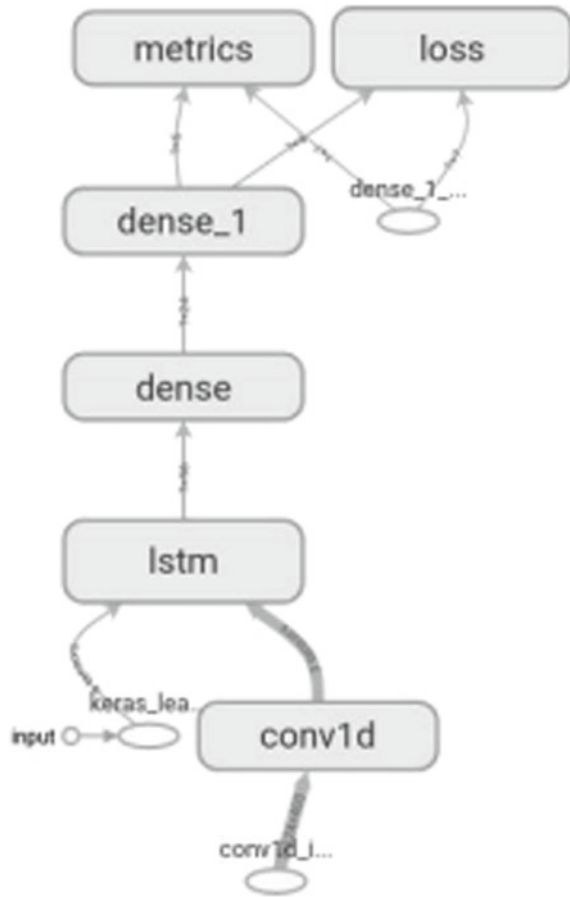
Fig. 2 CNN model architecture

faster in the absence of any negative values. The output vector shape of this layer is (24). It serves as an input to our output layer.

Output Dense Layer. For our output layer also, we used a fully connected dense layer with six nodes, with softmax activation function, which provides an output vector representing the probability likelihood of each of our six marks labels. Hence, while predicting, we can tell the marks as the one with the highest degree of probability. This layer generates an output vector of size (6).

The model was compiled using the Keras API compile function with the Adam optimizer as the optimizer function, sparse categorical cross-entropy for loss computation and accuracy as the evaluation metric of the model (Fig. 3).

Fig. 3 LSTM model architecture



6 Result and Analysis

Random forest classifier was used as a baseline reference model for the proposed deep learning models. The proposed prediction-based system gives us a predicted mark for a corresponding answer, derived from neural networks. The model was trained over 10 and 15 epochs for CNN and LSTM, respectively.

Cross-validation is a resampling procedure to validate the effectiveness of a developed model if a limited data source is available. We used cross-validation for testing the accuracy of the model by varying the training fraction from 0.2 to 0.8. Accuracy metric of Keras library was used for calculating the accuracy of the model. There is a list of Keras classification accuracy metrics like binary accuracy, categorical accuracy, sparse categorical accuracy, etc.; regardless of whether the problem is a binary or multi-class classification problem, we can specify the metric as 'accuracy' to report on accuracy. Total and count are two local variables created by accuracy

metric where total and count here represent the number of predicted marks that matched with the actual mark and total number of samples, respectively. Accuracy is finally calculated as the frequency of total by count.

For training purposes, we have used 10 epochs for CNN and 15 for LSTM. The accuracy rate per epoch is ascending along with the descending nature of epoch loss for both of them. The graphs of epoch accuracy and epoch loss for both CNN and LSTM are given in Figs. 4 and 5, respectively.

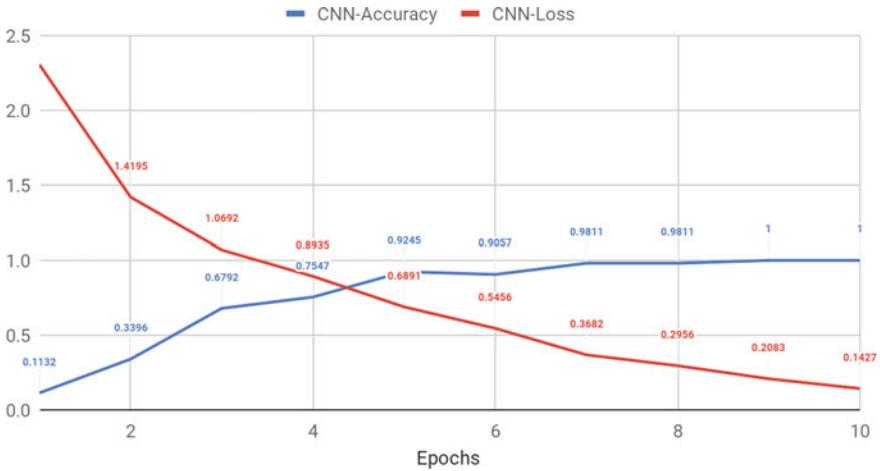


Fig. 4 CNN loss and accuracy graph

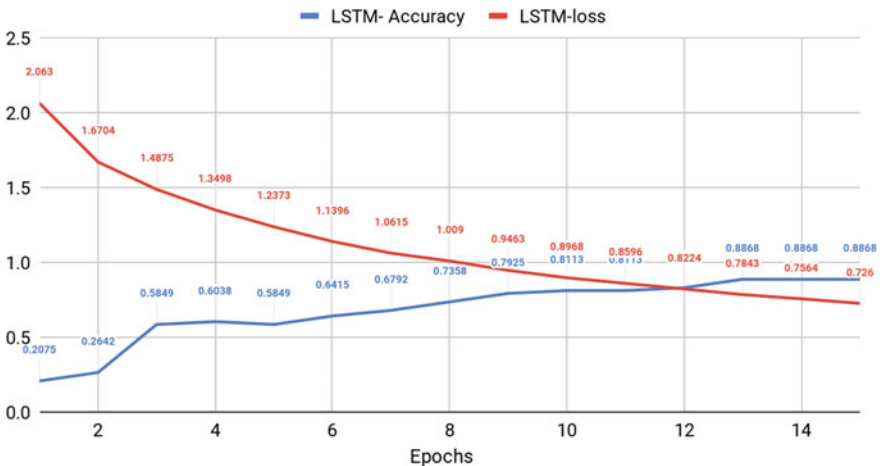


Fig. 5 LSTM loss and accuracy graph

Table 2 Accuracy of various implemented models at different training fractions

Training fraction	RFC	CNN	LSTM
0.2	35.185185	35.185185	45.185186
0.3	44.680851	44.822695	52.907801
0.4	39.024390	49.756098	51.951221
0.5	52.941176	56.274511	54.803923
0.6	51.851852	53.703705	54.444446
0.7	57.142857	54.285716	63.174605
0.8	42.857143	62.142858	74.761905

The models were trained and tested over 30 iterations to obtain a normalized accuracy. Table 2 depicts the comparison of the normalized accuracy levels at different train-test splits of the dataset.

The accuracy of the model at 80% of training data is 42.85% for the random forest classifier (RFC), while convolutional neural network (CNN) and long short-term memory (LSTM) show an accuracy of 62.14% and 74.76%, respectively. The corresponding graph for a better representation, understanding and easy comparison of accuracies is given in Fig. 6.

Clearly, the LSTM model has outperformed both the baseline RFC model and CNN model. Therefore, the LSTM model was chosen for ‘Descriptive Indic Answer Script Evaluation using Deep Learning.’

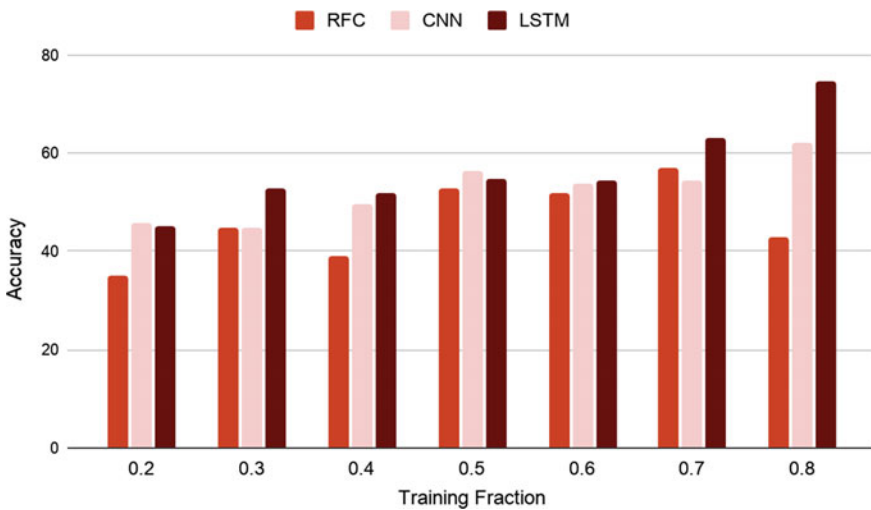


Fig. 6 Training fraction versus accuracy

7 Conclusions

A lot of time and concentration is required during manual answer sheet evaluation. The evaluator can commit mistakes as there are heaps of answer sheets to be evaluated. To mitigate these problems, ‘Descriptive Indic Answer Script Evaluation using Deep Learning’ was proposed. In this proposed system, we mainly focussed on Indic languages as almost negligible work has been done in these languages. For this system, both classical machine learning model and deep learning models were implemented. The highest accuracy was obtained from the LSTM model which was 74.76%. As there was no precedencing work exist in this field, thus the accuracy of 74.76% is quite acceptable. The accuracy obtained in the model has been compared with the existing works in English language. From the findings, it is evident that the proposed model works quite comparable to the existing systems in English language.

References

1. T.K. Landauer, Automatic essay assessment. *Assess. Educ. Princ. Policy Pract.* **10**(3), 295308 (2003)
2. T. Kakkonen, N. Myller, J. Timonen, E. Sutinen, Automatic essay grading with probabilistic latent semantic analysis, in *Proceedings of the 2nd Workshop on Building Educational Applications Using NLP Association Senapati et al., International Journal on Emerging Technologies* 11(x): xx-xx(2020) 6 for Computational Linguistics, pp. 29–36 (2005)
3. P. Patil, S. Patil, V. Miniyar, A. Bandal, Subjective answer evaluation using machine learning. *Int. J. Pure Appl. Math.* (2018)
4. P. Singh, S. Sheorain, S. Tomar, S. Sharma, N.K. Bansode, Descriptive answer evaluation. *Int. Res. J. Eng. Technol. (IRJET)* **05**(05) (2018)
5. N. George, P.J. Sijimol, S.M. Varghese, Grading descriptive answer scripts using deep learning. *Int. J. Innov. Technol. Exploring Eng. (IJITEE)* **8**(5) March, 2019, ISSN: 2278-3075
6. S.G. AlperKürşatUysal, The impact of preprocessing on text classification. *Inf. Process. Manage.* **50**(1), 104–112 (2014)
7. K. Weiss, T.M. Khoshgoftaar, D. Wang, A survey of transfer learning. *J. Big Data* **3**, 9 (2016)
8. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering* ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh. *Lecture Notes in Electrical Engineering*, vol. 661. Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-4692-1_58
9. S. Mandal, S. Biswas, V.E. Balas, R.N. Shaw, A. Ghosh, Motion prediction for autonomous vehicles from lyft dataset using deep learning, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, 30–31 Oct. 2020, pp. 768–773. <https://doi.org/10.1109/ICCCA49541.2020.9250790>
10. J. Howard, S. Ruder, Universal language model fine-tuning for text classification, in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers), vol. 1, pp. 328–339 (2018)
11. Y. Kim, *Convolutional Neural Networks for Sentence Classification*. New York University (2014). [arXiv:1408.5882v2](https://arxiv.org/abs/1408.5882v2) [cs.CL]
12. H. Kim, Y.-S. Jeong, Sentiment classification using convolutional neural networks. *Appl. Sci.* **9**:2347 (2019)

13. A. Sherstinsky, *Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network*. Elsevier Physica D: Nonlinear Phenomena journal, vol. 404, Special Issue on Machine Learning and Dynamical Systems (2020)
14. J.-H. Wang, T.-W. Liu, X. Luo, L. Wang, An LSTM approach to short text sentiment classification with word embeddings, in *The 2018 Conference on Computational Linguistics and Speech Processing ROCLING 2018*, pp. 214–223 (2018)

Blockchain Aided Predictive Time Series Analysis in Supply Chain System



G. A. Dhanush, Kiran S. Raj, and Priyanka Kumar

Abstract In the modern digital era, Artificial Intelligence (AI) and Blockchain Technology are the two most popular and promising research areas. AI helps in making intelligent decisions whereas blockchain technology automates the payment process by making it more secure, trusted and decentralized. Smart contracts help with interactions among participants over the network with no intermediate or a trusted third-party involvement. In this paper, we have considered a real-time application of tracking crops produced by farmers across different stages of the crops supply chain. We have proposed a secure and trusted supply chain model using blockchain technology and have done time series analysis for future prediction for the quantity of production of each crop in a season.

Keywords Supply chain · AI · Time series analysis · Prediction · Blockchain technology

1 Introduction

Supply chain is one of the well-known application areas of blockchain technology. It ensures end-to-end information flows, products, services and money. It also records each and every single asset throughout its flow on the supply chain, track the orders, receipts, and payments in a secured and transparent manner. In past few years, we have seen blockchain as an promising research area and its significant applications in healthcare, in financial services such as digital assets and payment systems [8, 12],

G. A. Dhanush · K. S. Raj · P. Kumar (✉)
Department of Computer Science and Engineering, Amrita School of Engineering,
Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: k_priyanka@cb.amrita.edu

G. A. Dhanush
e-mail: cb.en.u4cse17415@cb.students.amrita.edu

K. S. Raj
e-mail: cb.en.u4cse17430@cb.students.amrita.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_70

913

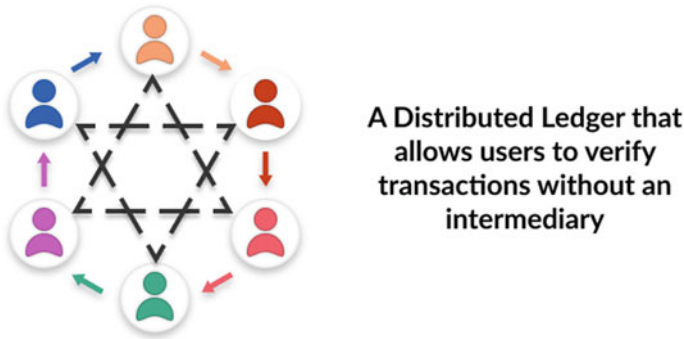


Fig. 1 Imagining blockchain (Decentralized, consensus controlled, tamper-proof, trustless public ledger of assets and transactions)

smart contracts [13], logistics [11] Internet of Things (IoT) [9, 16], and reputation systems [5, 10]. All this data resulting from supply chain is stored as raw data in the Blockchain. We can perform a variety of analysis on data of this magnitude to make predictions and aid the growth of companies participating in supply chain.

We live in India which is also called agricultural country. Lot of agriculture lands are there which is the largest sector for employment. India ranks second worldwide in farm outputs. As per 2016, agriculture employed 49.9% of the Indian workforce and contributed 16.9–17.9% to country's GDP [14, 15]. It would be of huge help to them if we are able to track their produce by storing the crop data in Blockchain. Since the data is recorded in Blockchain, the data becomes easier to track. At the end, we would like to fetch the data recorded in the blockchain and make useful time series analysis to know how much crop can be produced in a given season (Fig. 1).

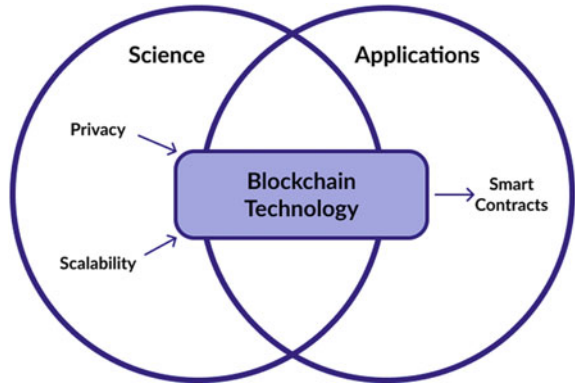
Road map. The paper is organized as follows. We describe literature survey on impact of blockchain in supply chain system in Sect. 2. In Sect. 3, we formally show the implementations of our proposed CropPrediChain model and have shown user validation and successfully tracking of crops at every stage of supply chain network. Section 4 describes time series analysis for future prediction for the quantity of production each crop in a season. Finally, we conclude in Sect. 5 (Fig. 2).

2 Background

2.1 Blockchain Technology and Smart Contract

Blockchain Technology has great impact across industries in 5–10 years. It is one of the most demanding technology for security, transparency and immutability purposes in every aspects of modern research domain. The following Hype Cycle provides an overview of how blockchain capabilities are evolving from a business perspective and

Fig. 2 Blockchain Technology



maturity across different industries [14]. Blockchain is also called distributed ledger. A ledger is a system of records for a business that records asset transfer between participants. It is immutable in nature which is achieved by hashing technique. Hash functions take an electronic record (such as a PDF file, a video, and an email) and produce a fixed-length output. If the information is changed in any way even a comma is changed in a 3000-page document, a different output value is produced. Hash is irreversible. An input gives a hash, but a hash cannot give the input (Fig. 3).

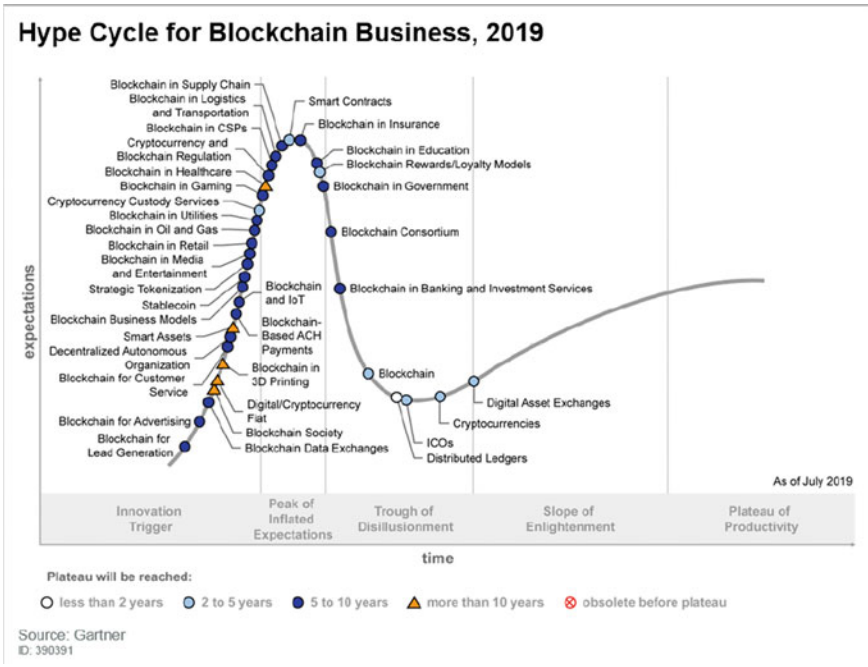
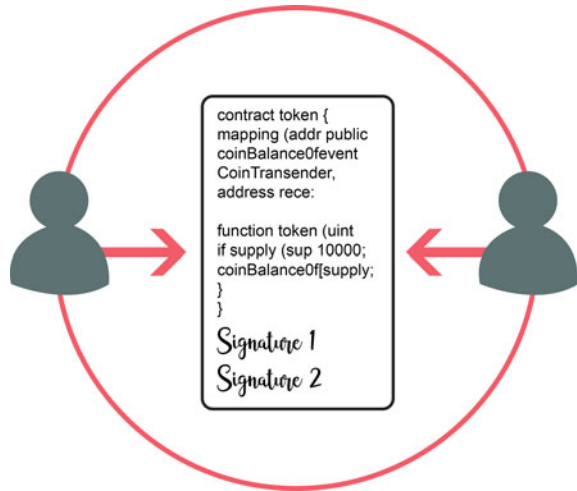


Fig. 3 Gartner Hype Cycle for Blockchain (September 2019) [4]

Fig. 4 Smart contract



Smart Contract is the digital agreement between transacting parties. This is a computer protocol intended to digitally facilitate, verify, or enforce the negotiation or performance of a contract. Smart contracts allow the performance of credible transactions without third parties. Contract is the part of public blockchain and is written in the form of code into this blockchain as depicted in the figure. Parties involved in the contract are anonymous. Contract execute itself when predefined condition is met (Fig. 4).

2.2 Limitations and Research Challenges in Supply Chain

Traditionally, supply chain was a linear process as seen in Fig. 5 starting from Tier N supplier, the product passes through different stages and reaches the dealer. A company starts with a small supply chain and as the business grows, many partners performing the same process come into the picture. It becomes difficult to keep track of the flow of product as many companies with similar roles come into the picture. Also, the communication between companies in one stage the supply chain to another company in another stage the supply chain becomes difficult as the chains become longer. All this could be mitigated with the use of blockchain. Blockchain helps in keeping track of all the processes that take place in the supply chain (Fig. 5).

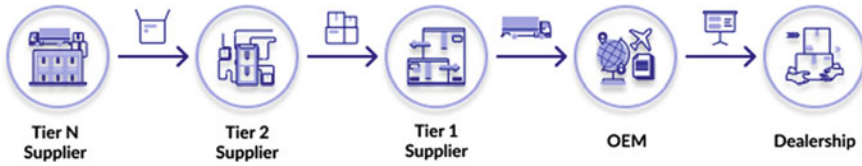
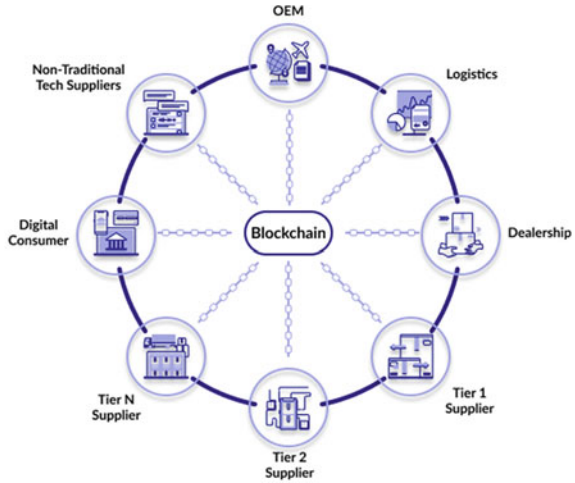


Fig. 5 Supply chain without Blockchain

2.3 Impact of Blockchain in Supply Chain

Latest blockchain technologies help to feed the data into the blockchain in real-time. Once fed, the data in the blockchain cannot be mutated. Hence, the integrity of the data is always preserved. Also, the origin of the data can be easily traced as all the record from the beginning of the process in the supply chain can be stored and retrieved. This makes blockchain the central mode of sharing and retrieving data for the companies involved in supply chain as shown in Fig. 7. When a company needs to remove redundant participants of its supply chain, it can easily query the blockchain to find out what companies work at a stage and filter out necessary companies. This reduces the cost incurred to the business in supply chain and makes it easier for companies to manage supply chain as it grows (Fig. 6).

Fig. 6 Supply chain with Blockchain



2.4 Technology Stack

To store crop details in Blockchain, we have used Ethereum [1] platform. This platform also provides support to write smart contracts. Hence, it becomes easier to model data and data structures and store the changes of these details and immutable transaction records in the blockchain. Solidity is the programming language used to write smart contracts and it is a language similar to JavaScript.

To bootstrap the process of setting up a development environment for writing smart contracts, we use Truffle [7] framework. This framework gives us a template to write smart contract and helps to deploy code in an Ethereum network. To test the smart contract code, we use a tool called Ganache [3]. It simulates Ethereum blockchain and helps us debug the smart contract that we have written.

Once the smart contract is deployed, it can be interfaced in JavaScript with the help of Web3js library. To build our frontend for our mobile application, we have used React Native [6] to create a functional interface. Using Web3 APIs, we interact with the smart contract deployed in Ethereum.

The data stored in blockchain is then fetched by a Flask [2] server written in Python. It pulls the data from blockchain periodically using Web3 library for Python. Entire time series algorithm is written in Python without use of any framework. The data fetched is fed into this time series algorithm and the predicted result is plotted as a graph as shown in Fig. 11.

3 A Blockchain Framework: CropPrediChain

Blockchain can be used to track the status of the crop produced by the farmer. By using blockchain, we ensure that the integrity of the data is not lost while remaining transparent to its stakeholders. The supply chain of the food product starts with raw material (crop) being produced by farmers. To store the information related to the crop produced by farmers, we create a ‘Farmer’ data structure with fields related to a crop. An application used by the farmer generates a unique ID for a crop. The farmer would enter the details of following fields along with this unique ID in the data structure:

- *FoodId*: Unique ID generated for a crop.
- *FarmerName*: Name of the farmer or the farm in which the crop is produced.
- *Location*: Location of the farmland.
- *CropName*: Name of the crop.
- *Quantity*: Weight of the crop produced (in kilograms).
- *Season*: Season in which the crop is grown.
- *FarmPrice*: Price at which the crop is sold (Fig. 7).

Once the farmer sells his product to an intermediary, he enters the selling price into the data structure. Now, depending on the crop and its utility, there could be

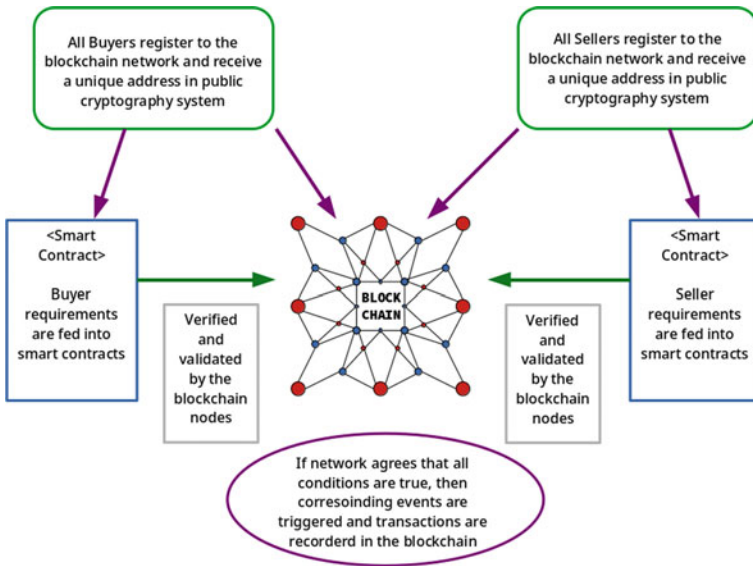


Fig. 7 A blockchain framework: CropPrediChain

different intermediaries in the supply chain. To be flexible in accommodating all the intermediaries, an ‘Intermediary’ data structure is created. It contains a stage and intermediary-type fields that tell in which stage of the supply chain the product is and what intermediary is handling the product, respectively. Crop’s unique ID is used to track a farmer’s produce. Following are the fields of intermediary data structure:

- *FoodId*: Unique IDs used to identify a crop used by the product.
- *Stage*: A number which tells at what stage of the supply chain the product is in.
- *IntermedType*: This field denotes the type of intermediary using the product.
- *IntermedName*: Name of the company or person handling the product.
- *ProdId*: If the raw material is transformed into a product, it gets a new unique product ID; Else, the field is left blank.
- *FoodProd*: Name of the product produced. If no product is produced, the field is left blank.
- *FoodProcQuantity*: Quantity of food procured from the previous stage.
- *FoodDistQuantity*: Quantity of food supplied to the next stage.
- *IntermedPrice*: Price at which the product is sold to the next stage.

After the food passes through a chain of intermediaries, it reaches the consumer through the retail store. At this stage, purchase made by each customer is kept track of. A data structure called ‘consumer’ is used to keep track of the data collected. Each transaction is given unique customer ID and the corresponding product id of the product purchased is recorded in this data structure. Following are the set of fields used in the data structure:

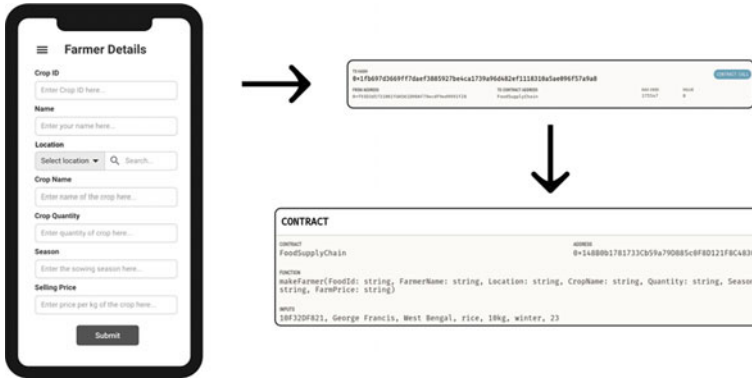


Fig. 8 Screenshot of application storing details in blockchain

- *CustomerId*: Unique ID is given to each transaction made by the customer.
- *ProdId*: Unique ID that denotes a product.
- *FoodQuantity*: Quantity of the product purchased.
- *RetailerName*: Name of the retailer selling the product to the consumer.
- *Location*: Location of the retailer.
- *ConsumerPrice*: Price at which the consumer purchases the product.

A unique ID for a consumer is generated for each transaction so that the user is not tracked by the products that he purchases. This way, we ensure that the product is tracked while ensuring the privacy of the customer.

After storing the data in the blockchain, important fields such as location, crop name, quantity, season, the price at which the crop is sold, the price at which the final product is retailed and where the product gets retailed are fetched and feed into the time series analysis algorithm to predict the price of the crop at different times in a year. With these predictions, a farmer could understand how his crop as a raw material gets distributed, find out potential regions where his crop can get sold and predict how much profit he could earn from selling the crops at these regions in advance (Fig. 8).

4 Time Series Analysis

4.1 Why Time Series?

To understand this, one must go to the roots of the problem that is currently being addressed. The farmers in India are being blinded by the fact of the price of the crops sold in the markets. They are given a mere penny amount compared to the colossal amounts by which they are sold at the market. The farmers are blinded and in turn,

end up in the darkest of ways to live. The main factor behind this is transparency, the farmers are not aware of the prices being sold in the end-markets and how much the middleman profit out of this.

There are many factors that affect the price of the crop. Since all the crops are seasonal in nature and are time dependent. The price of the crop in each month would fluctuate depending on the time of year. With the help of time series, one can predict what would be the price in the upcoming months depending on the data from previous months. This prediction can be used to alert the farmer regarding the sale of his crop in the end markets. Once the farmer is made aware, he can make necessary decisions regarding the same.

4.2 Recommendation System

With the power of Machine Learning and Associative-Based Learning, farmers can be given recommendations about what crops could be planted or sold at what point of time for better profits and efficiency. Taken into crop all the parameters needed for the growth of crops and the seasonality, once can give suitable recommendations regarding what crop would be suitable at what point of time.

4.3 Analysis

Auto-correlations are used to calculate the relations between observations as a function of time. Then statistical analysis such as the Dickey-Fuller test is used to test for the stationarity of the data. Stationarity is done to check whether the mean and variance of the distribution remain constant over a period. Such analysis if true is easy to model as the point of fluctuations happen over a pair of parameters such as mean and variance. Extrapolations and model smoothening is easier hence. The observations are decomposed preferably using seasonal decomposition to find out the trend, periodicity, and residue. With the help of smoothing techniques such as Exponential Smoothing, a priority weightage can be given to extract the trends for the future. The seasonality is deduced with a given window time frame and rolling window calculation. With the help of algorithms like Seasonal Auto-Regressive Integrated Moving Average, the expected extrapolation can be estimated. Using Reinforcement learning initiation, the best possible bound for each situation can be calculated and simultaneously updated.

4.3.1 Analysis of Time Series

The analysis of time series revolves around the concept of the correlations between the data points and the rate of change of the influx data. Time series is composed of

components through which the visualization of the information carried by the time series is inferred. When the time series is decomposed, we can infer the components as Cyclic, Trends, Cyclical and Irregularity.

4.3.2 Component Analysis

- Secular Trends—The skeleton report generates the increase and decrease in the flow of the time series graph. It generalizes whether the pattern of the series increases or decreases.
- Seasonal Variations—The influx of time series data is affected by certain factors that could change the course of the flow of the series data.
- Cyclical Variations—These types of variations usually refer to the periodicity of the data, how frequent the data repeats in cycle is usually covered by the Cyclical Variations.
- Irregularity—Certain variations in the time series data are influenced by unpredictable influences, there usually does not exist techniques to analyze the same. Could also mean as error or the delta variations.

4.3.3 Time Series Data Decomposition

$$\begin{aligned} \text{timeseries}(t) \rightarrow & \text{seasonal}(t) \times \text{trends}(t) \\ & \times \text{cyclical}(t) \times \text{irregularity}(t) \end{aligned}$$

4.3.4 Stationarity of Time Series

An essential component which comes into play when using time series analysis is stationarity of the process. Stationarity is reached when the mean and variance of the functions remain is not subject to change with respect to time. Being stationary reduces the mathematical complexity, makes it easier to predict and analyze the data points. Hence a time series data $f(t)$ is said to be strongly stationary when the data points remain independent of time for the whole set. A time series data $f(t)$ is said to be weakly stationary when the moments of data depend on the difference of time and not the actual time itself.

$$\{x(t) \rightarrow x_1, x_2, x_3, \dots, x_n\}$$

4.3.5 Trend Analysis

$$m(t) = \sum_{x=-y}^y \frac{x_{t+y}}{1 + 2t}$$

In the above equation, spatial invariance is given consideration here. The formula is used to decompose the given time series graph into a skeleton trend graph depicting the rise and falls in the graph. Here $m(t)$ is the trend analysis function, t is the temporal variable (time), $-y$ to $+y$ refers to the window size of analysis, $x(t)$ denotes the given time series graph.

After analyzing the data, the models can be trained using a LSTM or other RNN's which hold transition pattern memory, in order to get appropriate results for the same (Fig. 9).

1. X-axis: The temporal quantity (time in months).
2. Y-axis: The quantity of crops produced in the specific months (Fig. 10).

1. X-axis: The temporal quantity (time in months).
2. Y-axis: The quantity of crops produced in the specific months.
3. Predicted quantity (orange line), Actual quantity (blue line).
4. From the graph one can infer that, the prediction incorporates most of the attributes represented by the ground truth data. Hence, we can conclude that even for a small dataset, the accuracy is in par with the ground truth analysis, there adhering to the change/ fluctuations within the patterns of the time series (Fig. 11).

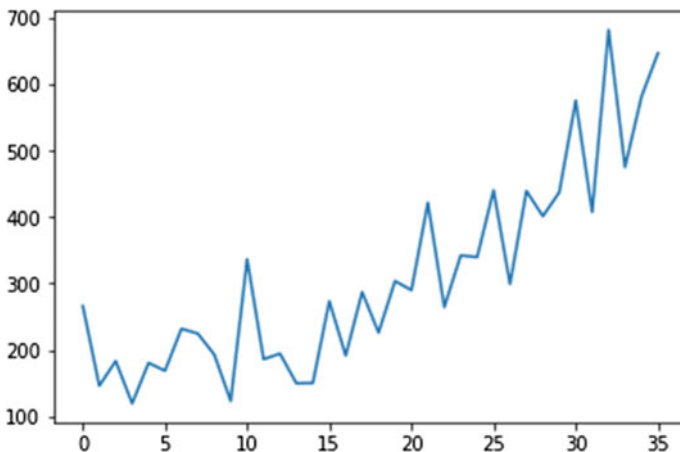


Fig. 9 Analysis of amount of crops produced over an year

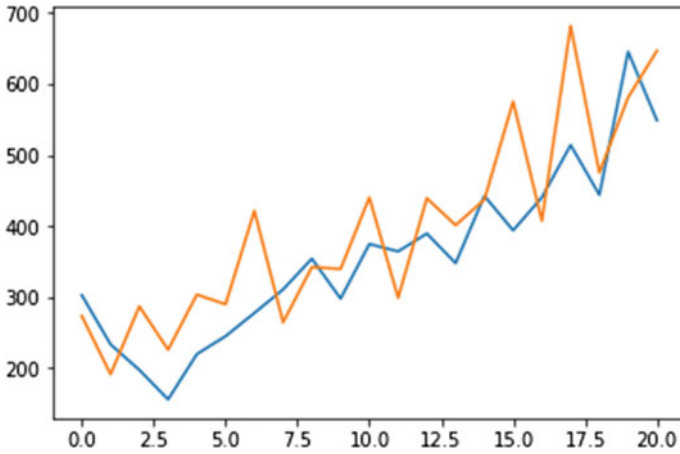


Fig. 10 Prediction of what the crops would be in upcoming months

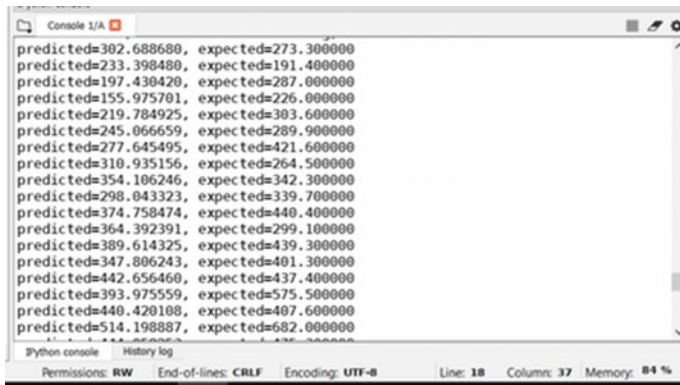


Fig. 11 Prediction of data points from time series

From the first row, we see the values of:

1. predicted = 302.688680
2. expected = 273.300.

These values denote the prediction values for the given dataset. The expected value is the actual value and the predicted is the predicted value from the model. Even for a very small dataset, the model accommodates most of the variations described the data in the trends.

5 Conclusion and Future Work

In this paper, we had explored the advantages of using block-chain in supply chain management. Additionally, we have taken a scenario of crop prediction and seen of data collected from different stages of crop supply chain can be stored in blockchain and later retrieved to perform analysis and make prediction on how much crop can be produced in a given season. This indicates that blockchain can be used as a central piece in supply chain to process data. Also, the data stored in the blockchain is also suitable to be processed by machine learning and deep learning algorithms and gain key insights from them. In future, we would like to focus and improve the security of the data stored, improve the performance of the analysis algorithm, increase the throughput of data and find other analysis techniques that might benefit from the data stored in the blockchain.

References

1. Ethereum (2019). <https://ethereum.org>
2. Flask (2019). <https://palletsprojects.com/p/flask>
3. Ganache (2019). <https://www.trufflesuite.com/ganache>
4. Gartner's hype cycle (2019). <https://www.gartner.com/en/newsroom/press-releases/2019-09-12-gartner-2019-hype-cycle-for-blockchain-business-shows>
5. India's biggest challenge in future farming (2019). <https://www.theindiaforum.in/article/india-s-biggest-challenge-future-farming>
6. React-native (2019). <https://reactnative.dev>
7. Truffle (2019). <https://www.trufflesuite.com>
8. A. Baier, Organic certification process. National Sustainable Agriculture Information Service (2005). <https://attra.ncat.org/attra-pub-summaries/?pub=163>, available online
9. M. Dobrovnik, D.M. Herold, F.E.K.S.: Blockchain for and in logistics: What to adopt and where to start. *Logistics* **2**, 18 (2018)
10. F.T.: An agri-food supply chain traceability system for china based on rfid & blockchain technology, in *Proceedings of the 2016 13th International Conference on Service Systems and Service Management (ICSSSM)*, pp. 1–6. Kunming, China (2016-06-24)
11. H.M. Kim, L.M.: Towards an ontology-driven blockchain design for supply chain provenance (2016). <https://dx.doi.org/10.2139/ssrn.2828369>, available online:
12. M. Maouchi, O. Ersoy, E.Z.: Trade: A Transparent, Decentralized Traceability System for the Supply Chain (2018). http://dx.doi.org/10.18420/blockchain2018_01, available online
13. G.W. Peters, E. Panayi, C.A.: Trends in crypto-currencies and blockchain technologies: A monetary theory and regulation perspective (2015). <http://dx.doi.org/10.2139/ssrn.2646618>, available online
14. M.S. Priyanka Kumar, To build scalable and portable blockchain application using docker, at soca 2019, 27th dec-29th dec 2019, nit. *Int. J. Comput. Netw. Commun. (IJCNC)*(ISSN) 0975-2293 (2019)
15. P. Kumar, G.A. Dhanush, S.D.N.A.S.S.S., An efficient and novel buyer and seller's distributed ledger based protocol using smart contracts, 11969, –349–363, (2020-01). https://doi.org/10.1007/978-3-030-36987-3_23
16. P/ Verhoeven, F. Sinn, H.T., Examples from blockchain implementations in logistics and supply chain management: exploring the mindful use of a new technology. *Logistics* **2**, 20 (2018)

Regional Language Code-Switching for Natural Language Understanding and Intelligent Digital Assistants



S. B. Rajeshwari and Jagadish S. Kallimani

Abstract Due to their widespread use around the globe, especially in a multilingual country like India, digital assistants are programmed to be capable of understanding speech and conversing for an array of languages. The objective is to understand user intent and respond with human expression, achieving a great degree of accuracy and efficiency. However, a multilingual society such as India has grown quite comfortable using a combination of two or more languages while communicating, commonly referenced as code-switching. Code-switching is not inherently taken care of by today's intelligent assistants. Due to the lack of standardized dataset, the presented work includes corpus generation of code-switching frequently asked queries was generated. It includes a study of intent classification of code-switching queries with various vectorizers and classifier combinations. This work also comprises a pipeline to identify the user intent and provide a desired response using keyword extraction and named entity recognition models on code-switching queries. Standardized code-switching metrics applied on the generated corpus to estimate the strength of code-switching yielded high intensity. The results of the intent classification study on code-switching queries obtained by evaluating the models on standardized classification metrics have yielded high accuracy. Named entity recognition in code-switching queries also have yielded high precision and recall scores.

Keywords Named entity recognition · Code-switching · Evaluation metrics · Intelligent assistants · Spoken dialects · *Hinglish* corpus

S. B. Rajeshwari
Department of Computer Science and Engineering, M S Ramaiah Institute of Technology,
Bangalore, India
e-mail: rajeshwari.sb@msrit.edu

S. B. Rajeshwari · J. S. Kallimani (✉)
Visvesvaraya Technological University, Belagavi, Karnataka, India
e-mail: jagadish.k@msrit.edu

1 Introduction

In this epoch of digital technology, the rise of intelligent assistants is eminent as they are powered by advancements in domains of artificial intelligence, machine language and natural language understanding. Individuals find using assistants convenient as all they have to do is make a query while the assistant undergoes the process of understanding the user intention and providing desired responses to the user. Currently, intelligent assistants are capable of handling varied languages and are very effective in providing services to users.

Considering linguistically diverse countries like India, Russia, Switzerland and many others that have multilingual speakers that converse in second language or third language too as a part of everyday life, it is natural for such speakers to have a tendency to switch languages or use multiple elements from different languages to convey their thoughts in a single context of conversation, which is called code-switching.

Therefore, though intelligent assistants are programmed to be capable of understanding human speech and conversing in an array of languages, code-switching is not inherently taken care of by today's intelligent assistants. Hence, speakers that replace certain subsets of a language's vocabulary with another language when conversing end up being misinterpreted by the assistant and hence fail to achieve desired response. Therefore, this project aims at providing facilitation to Indian multilingual speakers that query in code-switching requests of Hindi-English(*Hinglish*) and achieving desired response from the intelligent assistant.

2 Related Works

Aabi [1] presented his study on defining code-switching in the syntax of Arabic and French code-switching in Morocco. This paper provided a general introduction to code-switching, explains the need and scope of the topic which provided insights that code-switching was a trending topic in NLP research domains. Sarah Schulz et al. [2] also presented the various challenges of computational processing of code-switching, mentioning issues of language identification in code-switching texts, and also stating automatic language identification for the code mixed social media texts is a challenging task. A word level language identification experiment was performed by them using a trigram-based and conditional random field (CRF)-based model.

Pratik et al. [3] made use of a skip gram vectorizer and recurrent neural network to perform intent classification for task-oriented dialogue systems, over a corpus of code mixed Hindi-English queries, translated from the SNIPS NLU dataset. While this work is code-switching in nature, the corpus developed is not benchmarked over a valid set of metrics, as stated by Khanuja et al. [4]. It is also observed that the corpus consists solely of Hindi-English code-switching queries.

Simran Khanuja et al. presented an evaluation benchmark, GLUECoS, for code-switching languages, that spans several NLP tasks in English-Hindi including

language identification from text, POS tagging, MNER, sentiment analysis, question answering and a new task for code-switching, natural language interface. This paper was considered as the base paper for code-switching corpus generation and validating the corpus on the standardized code switch metrics mentioned. The datasets used in this project are benchmarked by the GLUCoS metrics and also incorporate both of Hindi-English and Kannada-English queries.

Lyu et al. [5] presented their work on speech recognition on code-switching among the Chinese dialects where an integrated approach was proposed to do automatic speech recognition on code-switching utterances, where speakers switch back and forth between at least two languages. Similar work was presented by Chan et al. [6] on automatic speech recognition of Cantonese-English code mixing utterances that described recent work on the development of a large vocabulary, speak-independent, continuous speech recognition system for Cantonese-English code mixing utterances. Further, Chen et al. [7] presented multi-task learning in deep neural networks for Mandarin-English code mixing speech recognition that proposed a multi-task learning deep neural networks approach that has been proven to be effective for acoustic modeling in speech recognition which is applied to Mandarin-English code mixing speech recognition. Long et al. [8] have also presented work in the same domain focusing on acoustic data augmentation for the Mandarin-English code-switching speech recognition task by proposing a code-switching acoustic event detection system based on the deep neural network to extract real code-switching speech segments automatically.

Mahata et al. [9] analyzed code-switching rules for English-Hindi code mixed text. They proposed an efficient and less resource intensive strategy for parsing and analyzing switching points in code mixed data. It explores the rules of code-switching in Hindi-English code mixed data. The work involves code mixed text extraction, translation of the extracted texts to its pure form, forming word pairs, annotation of these using parts of speech tags.

Jamatia et al. [10] presented work on POS tagging for code mixed English-Hindi Twitter and Facebook chat messages where the paper reports work on collecting and annotating code mixed English-Hindi social media text and experiments on automatic tagging of these corpora, using both a coarse-grained and a fine-grained POS tag set.

Bhargava et al. [11] show work on NER for code mixing in Indian languages using hybrid approach proposing an optimistic exponential type of ordered weighted averaging (OWA) operator as a hybrid recommender system.

Intent detection for code mix utterances in task-oriented dialogue systems was presented by Pratik Jayarao and Aman Srivastava which proposes an intent detection approach using supervised classifier models for code-switching data. They provided an extensive comparison between various vector models, encodings and classifiers based on their performance on code-switching data.

The textbook [12] is considered as a base for understanding various NLP tasks and it provided an insight into basics of POS tagging, context-free grammar, and sentence construction rules for imperative and interrogative sentences.

3 Considered Constraints

3.1 Identifying the Code-switching Regions

There are no fixed regions where code-switching tends to occur. Certain speakers have a tendency to code mix while conversing where they replace certain subsets of a language's vocabulary with another language in a single sentence while other speakers code switch languages after a few sentences or a paragraph. The speakers code switch according to their comfort, consciously or unconsciously, or to express particular ideas clearly and hence the switch can happen any time during the conversation [13].

3.2 Ambiguity in Dialects

Different dialects within the same language can lead to ambiguities in identifying regions of code-switching.

3.3 Unavailability of Code-switching FAQ Dataset

Non-availability of open-sourced code-switching FAQ datasets with labeled intents. Code-switching dataset for the purpose of intent classification and querying is not generalized; hence, it is not available for this application.

3.4 Utterance and Pronunciation of Non-english Words in Code-switching Text

The words in English are defined and are strictly followed in case of text and speech, while the same is not expected in regional languages which varies from dialect and utterance. These ambiguities also contribute to spelling errors when considered in the form of text.

With the above considerations, the presented work aims at providing facilitation to Indian multilingual speakers by handling vernacular code-switching requests made by users while conversing with intelligent assistants and providing desired responses to the users [14]. Additionally, it throws insights on performance studies of cross-lingual models used in the pipeline and includes generation of corpus of code-switching Hinglish frequently asked questions and its statistics.

4 Proposed Model

Due to the lack of standardized code-switching frequently asked question datasets involving Hindi-English switches, the project involves the creation of such a corpus. A corpus of about 1000 code-switching frequently asked questions was created that included code-switching queries on various domains such as:

- Delivery queries with queries on order delivery details, payment details and refund details.
- Insurance queries with queries related to health, travel and motor insurance.
- Aadhaar queries with queries related to E-aadhaar, updation of aadhaar card and other general queries.
- Medical queries with queries on symptoms, treatment and prevention of diseases.
- Find nearest queries with queries to find the nearest location.
- Booking queries which include queries for booking a hotel, flight, bus, train or a table at a restaurant and enquiry queries on vacation details.
- Reminder queries to note, set an alarm or reserve calendar slots (Table 1).

Since a new corpus was generated, code-switching statistics of the data in terms of standardized metrics for code-switching are used to validate code-switching in the corpus. Some of the standardized code-switching metrics used are:

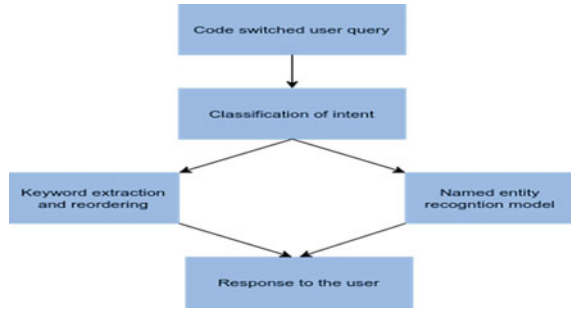
- **Multilingual Index (M-index):** A word-count-based measure quantifying the inequality of distribution of language tags in a corpus of at least two languages. The M-index is calculated as follows, where $k > 1$ is the total number of languages represented in the corpus, p_j is the total number of words in the language j over the total number of words in the corpus, and j ranges over the languages present in the corpus. The index is bounded between 0 (monolingual corpus) and 1 (each language in the corpus is represented by an equal number of tokens).

$$\text{M-Index} = 1 - \frac{\sum p_j^2}{(k - 1) \cdot p_j^2}$$

Table 1 Details of *Hinglish* corpus collection

Class	Number of samples	Number of unique tokens in Hindi	Number of unique tokens in English
Delivery queries	170	101	118
Insurance queries	155	74	181
Aadhaar queries	155	109	161
Medical queries	175	111	156
Find nearest queries	100	55	107
Reminder queries	100	84	138
Booking queries	150	160	199

Fig. 1 Overview of the proposed model



- **Language Entropy (LE):** The bits of information needed to describe the distribution of language tags. Language entropy is calculated as

$$LE = - \sum p_j \log^2(p_j)$$

and is bounded from below by 0 (representing a completely monolingual text) and bounded from above by $\log^2(k)$ which is the maximum entropy for a corpus with k languages (and, in such a case, each language is represented equally). In the case of two languages, the M-index and LE can be derived from one another (Fig. 1).

5 Implementation

The user provides a code-switching input either in the form of speech or text. If the query is in speech form, it is converted to text. If the query is in text form, the script of the text is checked. The entire text is to be transliterated to Roman script. This preprocessed text query is then passed to a supervised classifier to identify the type of intent. If the intent requires an action to be taken, then the query is passed to the code-switching named entity recognition model to identify entities such as date, location, restaurant names, cuisines, contact details, and then accordingly an action is performed by the assistant, whereas if an intent is one that requires information to be retrieved based on an user text input, it is passed through a sequence of steps that includes stop word removal, keyword extraction based on POS tags, language identification and translation to target language (English) and the reordering process to identify the words that contribute toward interpreting the intent of the user. Once the intent is identified from the search based code-switching query, a query is passed to Google and the relevant results are Web scraped and provided to the user (Table 2).

Table 2 Corpus statistics on code-switching metrics

Intent class	Multilingual index (M-index)	Language entropy (LE)
Delivery queries	0.926	0.972
Insurance queries	0.680	0.858
Aadhaar queries	0.988	0.995
Medical queries	0.979	0.992
Find nearest queries	0.984	0.994
Reminder queries	0.975	0.991
Booking queries	0.873	0.950

5.1 Intent Classification

The user can interact with a virtual assistant to perform a variety of different tasks, like querying for information, querying for local details, playing media, planning the day and many more. The act of correctly identifying what it is user wishes to accomplish with the help of a virtual assistant is defined as intent classification, with the action that the user wants the virtual assistant to perform being the intent.

It is essential to identify the type of the intent, so that the required action can be performed by the virtual assistant. A study was made on code-switching intent classification by using various supervised classification models with various vectorizing techniques to identify the efficient classification model. The different vectorizers used were the following:

Count vectorizer—The most straightforward vectorization method counts the number of times a token shows up in the document and uses this value as its weight. Since only the occurrence of the token matters, the language of the word and semantic meanings does not hold weightage in this technique.

TF-IDF—TF-IDF stands for “term frequency-inverse document frequency,” meaning the weight assigned to each token not only depends on its frequency in a document but also how recurrent that term is in the entire corpora. Again, neither the language of the word nor semantic meanings hold any weightage.

Word2Vec—Word2Vec is based on a distributional hypothesis where the context for each word is in its nearby words. Hence, by looking at its neighboring words, it can attempt to predict the target word. The skip gram architecture to generate Word2Vec was implemented where skip gram learns to predict the context words from a given word, in case where two words (one appearing infrequently and the other more frequently) are placed side-by-side, both will have the same treatment when it comes to minimizing loss since each word will be treated as both the target word and context word.

The various supervised classifier models implemented include Naive Bayes classification model, k-nearest neighbor classification model, random forest classifier, linear support vector classifier model, logistic classifier model and decision tree

classifier model. A statistical approach was preferred according to the availability of dataset size, since a smaller size of data could lead to over fitting during training neural networks like RNN, the results of BiLSTM were also compared with statistical models which provided proof to the assumption. Among various statistical models tested, the one with prominent results was given by the “linear support vector classifier.”

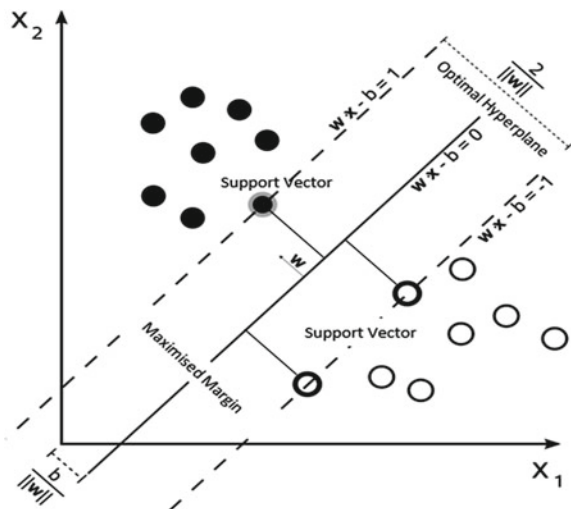
5.2 Linear Support Vector Classifier

A support vector machine (SVM) is a discriminative classifier formally defined by a separating hyperplane. In other words, given labeled training data (supervised learning), the algorithm outputs an optimal hyperplane which categorizes new examples into various classes. In two-dimensional space, this hyperplane is a line dividing a plane in two parts where in each class lay in either side.

An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. In addition to performing linear classification, SVMs can efficiently perform a non-linear classification, implicitly mapping their inputs into high-dimensional feature spaces. The support vector machine searches for the closest points as shown in Fig. 3, which are considered as the “support vectors” (Fig. 2).

According to the SVM algorithm, it finds the points closest to the line from both the classes. These points are called support vectors. Now, it computes the distance between the line and the support vectors. This distance is called the margin. The goal is to maximize the margin. The hyperplane for which the margin is maximum is the optimal hyperplane.

Fig. 2 Support vector classifier—hyperplane and support vectors



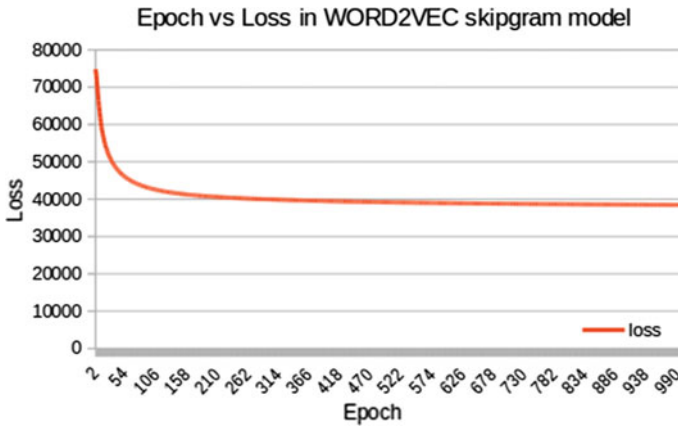


Fig. 3 Epoch versus loss in Word2Vec skip gram model

5.3 Stop Word Removal

Stop words are words in any language’s vocabulary that are commonly used and provide little to no value to the meaning of a sentence. The removal of stop words is preferred as it allows models to be trained on words that are truly contributive to the meaning of the sentence and also reduce dimensional space of the feature vectors generated from the count vectorizer. A stop word list is generated by sorting the frequency of words occurring in a given corpus in descending order, and deciding upon a cutoff below which words are not considered to be frequent. Different applications may or may not require the use of a stop list, and this dependency stems from a chance that a word or words in the list may provide some meaning to the target sentence or corpora. This project implements a stop word list consisting of Hindi, Kannada, and English words all transliterated to the *Devanagari* script.

5.4 Keyword Extraction Based on POS Tagging

Parts of speech (POS) tagging is the process of assigning a word class, such as a noun, verb, pronoun, etc., to a target word in a given sentence. These word classes, also called POS tags, give insight into the expected neighbors of a target word, and also relationships between words in the same sentence. They also act as features for a task called named entity extraction, which will be extrapolated upon in the coming sections.

There are three ways to train a POS tagger:

1. Using a HMM
2. Using a maximum entropy Markov model

3. Using a CRF.

The task in this project for code mixed POS tagging, the paper by Singh et al. is considered, who concluded that the CRF model provided the best results for POS tagging of short code mixed Hindi-English social media tweets. The advantage of bi-directionality is present in the CRF while absent in the rest of the proposed models. Thus, the same is implemented for POS tagging code mixed queries. In the task of intent classification, nouns, verbs, adjectives, and pronouns contribute the maximum meaning to a given sentence. Thus, words are extracted with these specific POS tags and discard the rest.

5.5 *Language Identification*

Language identification at the word level involves the tagging of each word in the input code mixed sentence with the language it belongs to. This is important as it helps to identify if the code mixed query is code mixed in the first place and also acts as a feature in assigning POS tags to each of the words in the input query. During the reordering of keywords obtained, it is required that every word in the code mixed query be in the same language. Thus, it is imperative that non-English words be translated to English, where in language identification identifies each word as an English, Hindi or Kannada word. This subtask also gives insight into the languages that are involved in the code mixed query, usually a combination English and Hindi, but can also be a combination of English and Kannada. Bhat et al. as a part of FIRE 2014 developed a language identification tool called *lit_cm* that could identify one of 7 different languages a word could belong to. The system made use of letter-based language models to classify a series of letters into one of 7 different labels, effectively becoming a sequence classification task. Google's language detects API, while patented has proved to have shown great accuracy in detecting the language of each word, and also has a well maintained open-source community. After comparing the pros and cons of the above two solutions, Google's language detect API was decided upon.

5.6 *Translating Keywords to Target Language*

For external search engines like Google, Bing, Microsoft, etc., it is observed that monolingual queries, English in particular, tend to give better search results, in terms of relevance to the query, as compared to code-switching queries. Along the same lines, there exist various monolingual task-oriented dialogue systems that understand monolingual queries much more accurately than code-switching ones. Word level translation was performed using the Google Translate API that has support for almost

all common vernacular languages. It was decided against training an in-house translator due to the lack of gold standard parallel data. The target language was chosen as English, due to its widespread understanding and data abundance for tackling further downstream tasks if needed, like sentence level machine translation, next word prediction, etc.

5.7 Reordering the Keywords

The extracted monolingual keywords, while capturing a majority of the information conveyed by the original code mixed query, does not retain order and consequently the relationships between each of the keywords. An imperative sentence is a sentence that resembles a request, or an order, or an intent to request something from someone or something.

Eg: Show me last week's homework.

An interrogative sentence is a sentence that resembles a question to something or someone.

Eg: What is the weather in Bangalore?

Based on the POS tags obtained for each of the keywords, rearrangement of list of keywords into one of the above sentence structures was constructed. This is done by making use of the sentence level construction rules for each of the above sentence structures, and expanding them till a single expression is obtained for both types of sentences, that retained the maximum information, had a grammatical structure, and was as simple as possible to attain the previous objectives. It then becomes a task of placing tag specific words into tag specific regions.

The expressions so obtained were:

Interrogative Sentences \Rightarrow pron_wh (pron)(adj)⁺ (noun)⁺

Imperative Sentences \Rightarrow verb (pron)(adj) * (noun)*

Pron_wh : Pronoun (Who, what, where, when, why, how)

Adj : Adjective

Verb : verb

Noun : noun

Pron : pronoun

+ : Regular Expression Semantic for 1 or more than 1 matches

5.8 *Web Scraping Queries*

Web scraping is a term used to describe the use of a program or algorithm to extract and process large amounts of data from the Web. Web scraping using Python is used to extract the data into a useful form that can be imported. After performing the above tasks that are translating entities to target language and reordering the keywords, the final query is passed to a Web scraping algorithm where it has the ability to scrape the Google search results for particular queries.

Web scraping algorithm uses Python beautiful soup module and requests library.

- Beautiful soup is used for pulling data out of HTML and XML files. It works with the parser to provide idiomatic ways of navigating, searching, and modifying the parse tree.
- Requests are used to send HTTP/1.1 requests extremely easily.

Using requests library, it is able to send HTTP request for Google search engine by using final query and get the search results in an object, then obtained result object is passed into beautiful soup object where it converts the HTML or XML parse tree, and from the parse tree, it is able to get the search data and use it for further tasks.

5.9 *Named Entity Recognition*

Named entity recognition is the task of assigning substrings of a given string into one of different predefined categories, like person names, organizations, locations, time expressions, quantities, monetary values, percentages, etc. NER systems allow the extraction of application/domain-specific keywords and values that can later be used as parameters for API and function calls. For example, if one wanted to lookup an individual in a database, based on the query provided to the bot, a “person” tag would have to be assigned to a word or series of words in the query. Then, the word/words that belong to the “person” category are parameters to a function that lookup a given name in the database.

NER systems can be built in one of three different ways:

- **Pattern-Based NER's:** Here, the entities to be extracted almost definitely have some sort of a structure that can be used to isolate certain parts of the query. Regular expressions and suffix/prefix-based heuristics are developed to capture these entities.
Eg: Dates, Phone Numbers, and Email IDs etc.
With more complicated named entities, regexes end up becoming very difficult to develop, as it is almost impossible to predict each and every variation in which a query can be asked by a user.
- **Dictionary-Based NER's:** named entities of the type Person Name, Restaurant Name, Location, etc., are easily detected with the help of an external lookup,

usually to a database or an API connected to a database. This is also referred to as ontology, or lexicon search.

The downfall of this approach lies in the fact that an absence of a named entity in an entity-specific lexicon can lead to misclassifications.

- **Context-Based NER's:** ML systems for NER can be developed with a bidirectional sequence classifier such as a CRF or HMM, that can use words around a target word, to determine if it belongs to one of the many named entities.

These models require the use of annotated data, with entities if present in a query, being correctly labeled.

Bi-directionality in this case allows the model to learn what comes before and after a given target entity word or string.

Open-source ML-based named entity extractors include Stanford Named Entity Recognizer (SNER), Spacy, NLTK, etc. The work heavily makes use of pattern and dictionary-based NER systems, with the addition of word/character level heuristics for dealing with code mixed queries. The NER system developed for this project is one of the first for code mixed queries, capable of identifying over 7 different named entities namely Location, Hotel Name, Number of People, Restaurant Name, Date, Time, Phone Number, Email ID and Activity.

6 Testing

6.1 *Word2Vec Skip Gram Model Loss Per Epochs*

The overall loss of Word2Vec skip gram model was computed after finishing each training sample according to the loss function. The loss function comprises two parts. The first part is the negative of the sum for all the elements in the output layer (before softmax). The second part takes the number of the context words and multiplies the log of sum for all elements (after exponential) in the output layer (Table 3).

6.2 *Eyeball Method of Testing for Word2Vec Skip Gram Model*

One way to evaluate the word2vec model is to develop a “ground truth” set of words. Ground truth will represent words that should ideally be closest together in vector space. For example, if the corpus is related to code-switching medical queries, the vectors for “treatment” and “ilaj” should have the smallest Euclidean distance or largest cosine similarity. Few test cases using eyeball methods are shown in Table 4.

Table 3 Loss versus epoch of Word2Vec model

Epoch	Loss
0	74,823.01
100	42,615.75
200	40,661.27
300	39,915.00
400	39,466.18
500	39,160.74
600	38,939.38
700	38,776.58
800	38,652.04
900	38,552.62
1000	38,471.69

Table 4 Test cases for Word2Vec skip gram model validation

Word 1	Word 2	Cosine similarity index	Hit
Vehicles	Gaadi	0.940	HIT
Treatment	Ilaaj	0.950	HIT
Food	Khaana	0.880	HIT
Health	Svaasthy	0.724	MISS
Insurance	Beema	0.884	HIT
Effects	Asar	0.945	HIT
Symptoms	Lakshan	0.919	HIT
Fees	Shulk	0.935	HIT
Letter	Patr	0.948	HIT
Nearest	Paas	0.761	MISS

6.3 Classification Metrics of Intent Classification

A study was made on code-switching intent classification by using various supervised classification models with various vectorizing techniques to identify the efficient classification model. The study was validated on standard classifier metrics including accuracy, precision, recall, F1-score and support.

Accuracy

Accuracy in classification problems is the number of correct predictions made by the model over all kinds of predictions made (Table 5).

$$\text{Accuracy} = (TP + FP) / (TP + FP + FN + PN)$$

Table 5 Comparison of accuracy of the classifier models

Classifier	Count vectorizer	TF-IDF	Word2Vec
Naive Bayes classifier	0.91265	0.88253	0.85240
K-nearest neighbor classifier	0.84036	0.85240	0.65060
C-support vector classifier	0.09036	0.09036	0.09036
Random forest classifier	0.89759	0.84939	0.89457
Linear support vector classifier	0.95180	0.93975	0.84638
Logistic regression classifier	0.94879	0.92771	0.44879
Decision tree	0.90662	0.82831	0.91265

Precision

Precision is the fraction of predicted positive events that are actually positive (Table 6).

$$\text{Precision} = \text{(TP)} / \text{(TP + FP)}$$

Recall

Recall (also known as sensitivity) is the fraction of positive events that are predicted correctly (Table 7).

$$\text{Recall} = \text{(TP)} / \text{(TP + FN)}$$

F1-score

The F1-score is the harmonic mean of recall and precision, with a higher score as a better model (Table 8).

$$\text{F1-Score} = \text{(2 * Precision * Recall)} / \text{(Precision + Recall)}$$

Support

Table 6 Comparison of precision of the classifier models

Classifier	Count vectorizer	TF-IDF	Word2Vec
Naive Bayes classifier	0.92	0.88	0.87
K-nearest neighbor classifier	0.88	0.86	0.77
C-support vector classifier	0.01	0.01	0.01
Random forest classifier	0.91	0.87	0.91
Linear support vector classifier	0.96	0.95	0.87
Logistic regression classifier	0.95	0.93	0.67
Decision tree	0.92	0.84	0.93

Table 7 Comparison of recall of the classifier models

Classifier	Count vectorizer	TF-IDF	Word2Vec
Naive Bayes classifier	0.91	0.88	0.85
K-nearest neighbor classifier	0.84	0.85	0.65
C-support vector classifier	0.09	0.09	0.09
Random forest classifier	0.90	0.85	0.89
Linear support vector classifier	0.95	0.94	0.85
Logistic regression classifier	0.95	0.93	0.45
Decision tree	0.91	0.83	0.91

Table 8 Comparison of F1-score of the classifier models

Classifier	Count vectorizer	TF-IDF	Word2Vec
Naive Bayes classifier	0.91	0.88	0.85
K-nearest neighbor classifier	0.84	0.85	0.67
C-support vector classifier	0.01	0.01	0.01
Random forest classifier	0.90	0.85	0.89
Linear support vector classifier	0.95	0.94	0.85
Logistic regression classifier	0.95	0.93	0.44
Decision tree	0.91	0.83	0.91

Table 9 Comparison of support of the classifier models

Classifier	Count vectorizer	TF-IDF	Word2Vec
Naive Bayes classifier	332	332	332
K-nearest neighbor classifier	332	332	332
C-support vector classifier	30	332	30
Random forest classifier	298	332	332
Linear support vector classifier	316	332	332
Logistic regression classifier	315	332	332
Decision tree	332	332	303

The support is the number of samples of the true response that lie in that class (Table 9).

6.4 Classification Metrics for NER

Performance of NER for each action-based intent, calculated by counting the total number of true positives and negatives, false positives and negatives for each entity

Table 10 NER metrics with respect to intent

Intent	Precision	Recall	F1-score
Hotel booking	96	71.2	83.177
Restaurant booking	100	60	75
Travel booking	100	88.88	94.12
Reminder	91	82	86.15

Table 11 NER metrics with respect to entity

Entity	Precision	Recall	F1-score
Date	100	87	93
Time	100	84	91.52
Hotel name	96	55	70
Location	100	88.88	94.12
Restaurant name	100	65.1	78
Activity	91	87.5	89.21
Number of people	100	100	100

associated with each intent, on a token basis. Precision and recall are calculated on the number of false/true positives and negatives. This is done on each and every intent that makes use of the NER workflow and is calculated on the basis of only intent specific entities, rather than all of the entities (Table 10).

Performance of NER, calculated by counting the total number of false positives and negatives and true positives and negatives, over each and every query that makes use of a given entity, on a token basis (Table 11).

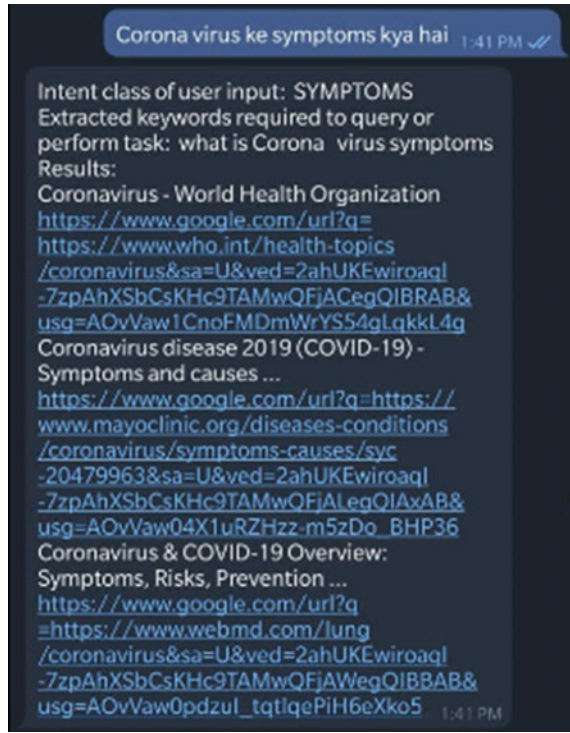
Samples of test case queries in code-switching made by the user and the assistant's responses are shared (Figs. 4, 5, 6, 7 and 8).

- Currently, the corpus generated has a high multilingual index for all the classes except booking and insurance classes while language entropy is high for all classes other than insurance class.
- The study on intent classification shows that linear support vector classification performs better than all the other classifiers when used along with either of count vectors or TF_IDF vectors with accuracy of 95% and 93%, respectively. The precision score and recall score when linear support vector classifier was used with count vector and TF_IDF was 96% and 95%.
- The responses provided by the assistant to the code-switching queries were observed to be relevant and pertinent.

7 Conclusion

This work created a robust interacting intelligent assistant which could respond to Hindi and English code-switching input queries from the user. The pipeline therefore

Fig. 4 Test case of code-switching query 1

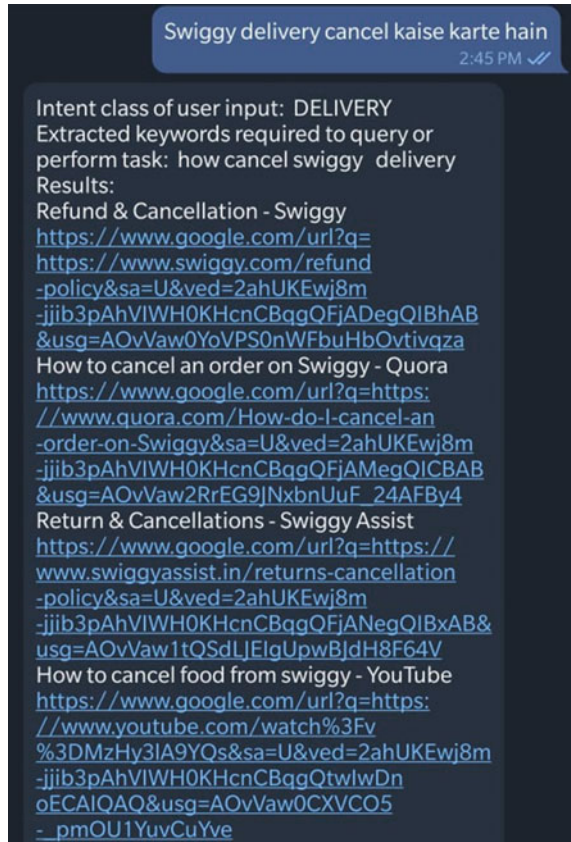


involved working with various natural language processing tasks on code-switching data including speech to text translation, intent classification, parts of speech tagging, named entity recognition, keyword extraction, and reordering words to meaningful sentences. A code-switching corpus was also generated with code-switching queries of frequently asked questions which had high strength when tested on standard code-switching metrics.

8 Future Tasks

The use of a speech to text SDK for automatic speech recognition tasks while convenient, does not incorporate strategies to enhance code mixed speech transcription. Therefore, a speech to text pipeline is currently being developed to accommodate for English-Hindi and English-Kannada code mixed speech input. Due to monolingual speech to text models being available for all three of the above languages, a custom speech to text architecture that can make use of these monolingual speech to text models at a word level will allow us to build robust English-Hindi and English-Kannada speech recognition systems.

Fig. 5 Test case of code-switching query 2



Pretrained multilingual supervised word embeddings (MUSE) published Facebook can represent words of two or more than two different languages in the same embedding space. That is, words that have some sort of connection or similarity in different languages are present near one and other. Along the same lines, multilingual bidirectional encoder representations from transformers (mBERT) published by Google also possess the same capabilities but have pretrained models that are used extensively in both industry and research. Making use of these pretrained models and fine-tuning with our dataset can boost accuracies and at the same time extend to other regional languages like Tamil and Telugu.

Fig. 6 Test case of code-switching query 3

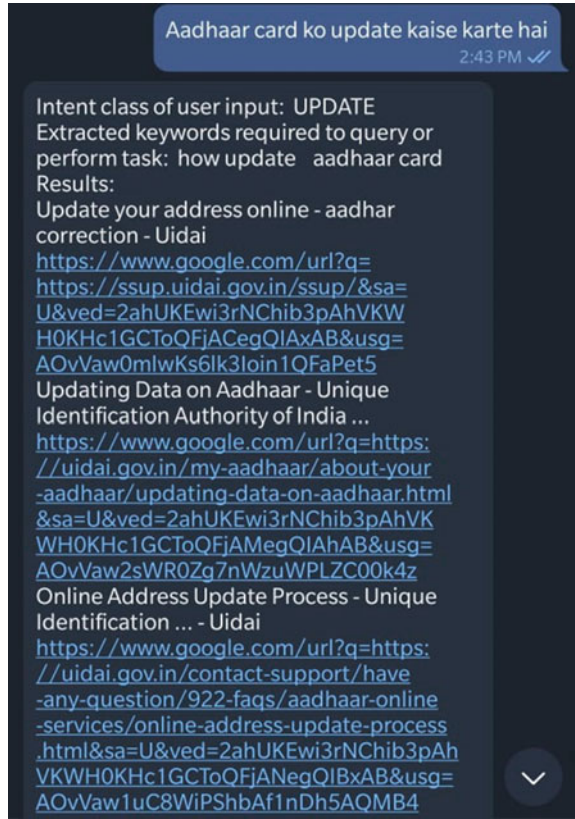


Fig. 7 Test case of code-switching query 4

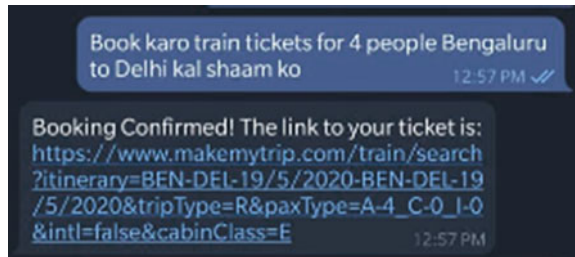
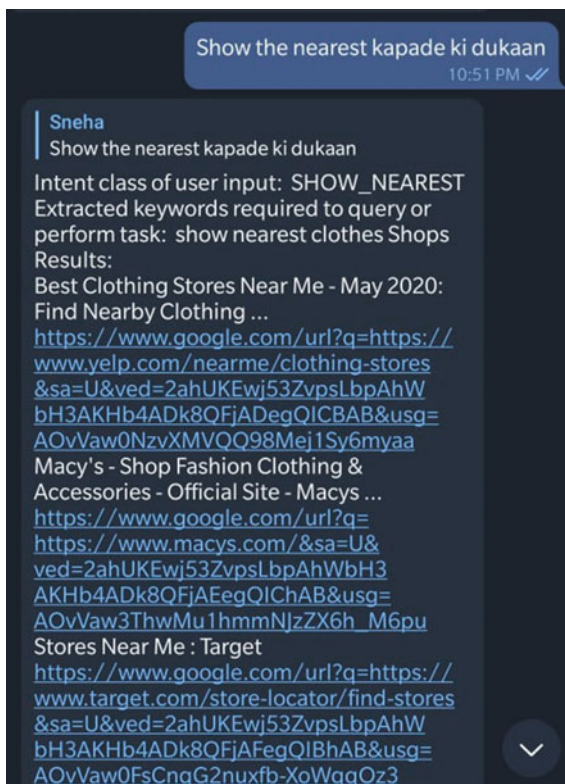


Fig. 8 Test case of code-switching query 5



References

1. M. Aabi, *Defining Code Switching in the Syntax of Arabic and French Code Switching in Morocco* (Springer, 2020)
2. S. Schulz, N.T. Vu, Challenges of computational processing of code switching, *ACI Anthol.* (2016)
3. P. Jayarao, A. Srivastava, Intent detection for code mix utterances in task oriented dialogue systems, in *Third International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT)* (Dec 2018)
4. S. Khanuja, S. Dandapat, A. Srinivasan, S. Sitaram, M. Choudhury, *GLUECoS* (Cornell University, An Evaluation Benchmark for Code Switched NLP, 2004).
5. D.-C. Lyu, R.Y. Lyu, Speech recognition on code switching among the chinese dialects, in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings* (May 2006)
6. J. Y. C. Chen, P. C. Ching, Automatic speech recognition of cantonese-english code mixed utterances, in *2006 Ninth International Conference on Spoken Language Processing* (Sept 2006)
7. M. Chen, J. Pan, Multi-task learning in deep neural networks for mandarin-english code mixing speech recognition. *IEICE Trans. Inf. Syst.* (2016)
8. Y. Long, Y. Li, Acoustic data augmentation for mandarin-english code switching speech recognition. *Appl. Acoust.* (2020)

9. S.K. Mahata, S. Makhija, *Analyzing Code Switching Rules for English-Hindi Code Mixed Text, Emerging Technologies in Modeling and Graphics* (Springer, 2019)
10. A. Jamatia, B. Gamback, A das, POS tagging for code mixed english-hindi twitter and facebook chat messages, in *10th Recent Advances of Natural Language Processing (RANLP)* (Bulgaria, Sept 2015)
11. R. Bhargava, B.V. Tadikonda, Named entity recognition for code mixing in Indian languages using hybrid approach, *CMEEIL Fire* (2016)
12. D. Jurafsky, J.H. Martin, Constituency grammars, speech and language processing: an introduction to natural language processing, *Computational Linguistics and Speech Recognition* (PH Publishing, 2019)
13. X. Li, Y.Y. Wang, D. Shen, A. Acero, Learning with a Click graph for query intent classification, *ACM Trans. Inf. Syst.* **28**, 3, Article 12, June 2010
14. K. Singh, I. Sen, P. Kumaraguru, A twitter corpus for hindi-english code mixed POS tagging. *Association of Computational Linguistics (ACL)* (2018)

Dimensionality Reduction for Face Recognition Using Principal Component Analysis Based Big Bang–Big Crunch Optimization Algorithm



Supreet Grewal and C. Rama Krishna

Abstract Dimensionality reduction is a significant step for various multimedia-based applications, where transmission bandwidth and storage space requirements are prodigious. In this paper, a new hybrid principal component analysis (PCA) based technique in combination with big bang–big crunch (BB-BC) optimization algorithm has been proposed. In this approach, PCA method is used for feature extraction and subsequently BB-BC algorithm performs feature selection. The purpose of this hybrid approach is to investigate the potential of dimensionality reduction, while retaining the maximum information. The recognition rate is considered to be the key performance indicator. The implementation of this method has been carried out using Olivetti Research Laboratory (ORL) face database. In the analysed scenario, this approach outperforms solely applied PCA technique with identical number of features selected.

Keywords Principal component analysis · Big bang–big crunch optimization algorithm · Soft computing

1 Introduction

With the surge in the use of the multimedia-based applications, high-speed Internet connectivity is imperative. Consequently, it results in the hike of bandwidth as well as capacity of server storage systems. It is evident from recent studies that approximately 90% of the Internet's traffic is occupied by multimedia-based data [1]. In pursuit to fulfil such obligations, dimensionality reduction becomes a critical phase. Its utilization is in the areas of distributed and network-based pattern recognition, multimedia processing and data mining applications. Many researches are dedicated in the direction to optimize dimensionality reduction of multimedia-based data. It

S. Grewal (✉)
Computational Intelligence Research Group Labs, New Delhi, India

C. Rama Krishna
National Institute of Technical Teachers Training and Research, Chandigarh, India

has provided an aid in storage capacity of server systems as well as higher Internet speed.

For dimensionality reduction, PCA is one of the prominent candidate. It is a statistical method which is widely used for data analysis, especially in the field of digital signal processing and neural computing. It aims to reduce dimensions of the input data by retaining relevant principal components in such a way that the maximum variance can be preserved. The literature includes application of PCA for image processing and face recognition. In a study, image compression was performed using neural networks and PCA method. Then, results based on total sum square error (TSSE) showed that PCA is preferable statistical tool than neural network technique to produce reduced dimensions to a diagnostic network [2]. Additionally, PCA is also deployed in joint photographic experts group 2000 (JPEG2000) to provide spectral decorrelation as well as spectral dimensionality reduction. The author realized PCA-based coder scheme, which is compared to JPEG2000 approach. The results suggested that PCA-based coder scheme performs the best [3]. In another study, PCA outruns the JPEG compression standard technique for compressing the medical images [4]. A newly suggested self-organizing map (SOM)-based technique was applied for dimensionality reduction in face recognition. It also uses PCA subsequently [5]. Out of all statistical and soft computing tools, PCA had proven to be the best resulting method for dimensionality reduction of image-type data.

PCA has been developed based on Eigen decomposition (ED) method, matrix theory of singular decomposition method (SVD) and Hebbian neural networks. These methods for dimensionality reduction had been used in face recognition as an area of application. ED and SVD are known as batch computing methods, while Hebbian neural network is adaptive soft computing method. The three methods were compared in terms of mean square error (MSE). In the literature, the performance of ED and Hebbian neural network have been identical and SVD has been the worst performer when only one image per class has considered. However, when the number of images per class was incremented, ED performed better than the other two [6].

The novelty of this paper is based on the improvements in recognition rate, when the hybrid method, namely PCA-based BB-BC optimization algorithm, is applied to ORL face database. It also provides a comparison between solely applied PCA and this hybrid method. Moreover, maximum recognition rate is achieved by including less number of features as compared to PCA single-handedly.

Hereinafter, the paper is organized as follows: in Sect. 2, the methodology to apply this hybrid approach is put forward, where PCA by ED and BB-BC algorithm are described. Afterwards, the results are computed and compared with hybrid approach and single-handedly applied PCA technique in Sect. 3. Ultimately, the conclusion remarks are presented in Sect. 4.

2 Methodology

Dimensionality reduction aims to reduce the dimensions without causing major degradation to the quality of images in the database. This involves two sequential steps, which are feature extraction and, subsequently, feature selection. Feature extraction is the process of transforming data from one coordinate system to another coordinate system. There are linear and nonlinear transformation techniques for feature extraction. In this paper, a linear transformation technique PCA is applied for feature extraction. ED method is used to perform PCA as it has been proved to be advantageous for dimensionality reduction. Afterwards, the feature selection aims to select an optimal subset of features from an extracted set of features. It helps to reduce the feature space, while retaining maximum information to reproduce original data. Optimal subset of features can be attained by using statistical tools [7, 8] as well as soft computing optimization techniques. In this work, BB-BC algorithm is proposed for feature selection. Finally, the comparison of two techniques would measure the performance.

In this work, the proposed hybrid approach makes application of PCA-based BB-BC algorithm for dimensionality reduction on ORL face image database. First step involves the division of input face image data into two data sets, which are training data set and testing data set. Afterwards, principal components are computed using PCA as per algorithm discussed in Sect. 2.1. To prioritize significant features, the computed principal components are arranged in decreasing order. All the features are extracted till the above step. Threshold value decides the point above which all the features are to be retained. The calculated value of threshold is thirty percent. BB-BC algorithm, as discussed in Sect. 2.2, is applied for selection of a subset of features from the set of extracted features aiming to maximize recognition rate. The testing set images and reconstructed images are compared on the basis of recognition rate. Image classification is realized using Euclidean norm as a distance measure.

2.1 PCA by Eigen Decomposition

PCA is optimal linear transformation method, used for feature extraction [9]. It transforms a number of correlated variables into a smaller number of uncorrelated variables. It aims to reduce mean square error in order to have candidate placed at first position with maximum variance count, and last position is taken by the candidate with minimum variance. For face recognition, Eigen face technique was presented by Turk and Pentland [10], which was based on ED for PCA. It is the standard and most extensively used appearance-based method that is used to implement PCA. The set of relevant features of face images are the principal components (PC) or Eigen faces. The procedure of proposed approach operates with the help of Eigen face technique, which is illustrated below:

Step 1 Get input data as face images.

Read the input face images as a matrix. Considering p as the number of images in a set and x_i represents an image, where $i = [1, 2, 3 \dots p]$.

Step 2 Subtract the mean.

Subtract the mean \bar{x} from each of data dimension to obtain $x_i - \bar{x}$.

Step 3 Calculation of covariance matrix.

Covariance matrix C_i is computed, which characterizes the relationship between variables of input matrix. It is given by

$$C_i = (x_i - \bar{x})(x_i - \bar{x})^T, \tag{1}$$

where T represents the transpose of the matrix.

Step 4 Compute the eigenvalues and corresponding eigenvectors of covariance matrix.

Eigenvectors and corresponding eigenvalues (features) of covariance matrix C_i is calculated in such a way that.

$$C_i V = A V. \tag{2}$$

where $A = \text{diag}(\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_p)$, a diagonal matrix defined by eigenvalues, λ_i of matrix C_i and $V = (V_1, V_2, V_3, \dots, V_p)$ are associated eigenvectors.

Step 5 Choosing components and forming feature vectors.

The eigenvalues and their corresponding eigenvectors are arranged in descending order. PCs are ordered according to their significance. Subsequently, first k number of eigenvectors is selected such that $k \leq p$ and discard left $p-k$ eigenvectors to find data in new directions. Here, k is known as knee point or threshold value.

Step 6 Deriving the new data set.

Finally, transpose of the formed feature vector is multiplied by left of the original data set transposed.

$$F = M \times N, \tag{3}$$

where M is the transposed eigenvector matrix and N is the transposed mean adjusted data. The threshold value, also known as knee point, is the spot above which all the values are considered to be significant, as the maximum variance lies in those values. It is calculated to extract the relevant features from complete set of features or eigenvectors. It can be calculated by plotting the graph between eigenvalues and the number of eigenvectors. The point where curve starts to saturate is considered to be the knee point.

2.2 Big Bang–Big Crunch Algorithm

The BB-BC algorithm, proposed by Erol and Eksin, is an evolutionary optimization method, inspired from one of the cosmological theories of universe [11]. This algorithm can be utilized for various applications to solve nondeterministic polynomial hard problems as it is capable of quick convergence. Here, extracted feature set from PCA method is input to BB-BC algorithm. The algorithm is as follows:

1. Initialize by including features from search space, into a set randomly. Population is the total number of features taken from the input extracted feature set, which is analogues to search space.
2. Calculate the recognition rate for all candidate solutions.
3. The best solution for the new search, which is centre of mass, is calculated using below equation. Best solution is a set of random features giving maximum recognition rate.

$$X_{CM} = \frac{\sum_{i=0}^N \left(\left(\frac{1}{f^{(i)}} \right) y_c \right)}{\sum_{i=0}^N \left(\frac{1}{f^{(i)}} \right)}, \quad (4)$$

where X_{CM} is the index of the single representative point attained and y_c is the distance of point from the centre. The fitness function value is denoted by $f^{(i)}$, which is recognition rate of feature i and N is the feature space size.

4. Calculate new feature set around the best fit candidate by adding or subtracting a normal random vector, whose value decreases as the iterations elapse. It can be formulated as:

$$X_n = X_{CM} + \frac{LR}{k} \quad (5)$$

where X_{CM} is best fit candidate, L represents parameter's upper bound, R is the normally distributed random number and k is iteration index. Then, the new point X_n which is upper and lower bounded, is calculated using the above equation.

5. Loop to Step 2 until maximum recognition rate reached to its convergence point.

3 Results

ORL face database constitutes 40 classes, where each class has 10 images. Each image has 92×112 resolution. For computational simplicity, only 20 classes with 10 different images of a person are considered. The resolution of these images is set to be 60×60 . From the face image matrix of 200 images, 40 images are included in training set that is 2 images per class. The rest 160 images are included in testing

Table 1 Recognition rate for the different numbers of selected features using two methods

Number of selected features	Recognition rate (%)	
	PCA	PCA-based BB-BC
4	78.75	87.520
5	78.75	90.625
6	81.87	91.875
7	83.75	91.875
8	86.25	91.875
9	83.75	91.875

set, i.e. 8 images per class. PCA is applied to the training set, and knee point is computed to be 30%. Hence, first 12 features out of 40 are retained. Afterwards, BB-BC algorithm is applied to select an optimal subset of features from total set of 12 features. The testing set is used to calculate recognition rate.

Recognition rate for the varied number of features included in the subset, for both PCA and BB-BC algorithm, is shown in Table 1. Recognition rate is plotted against the number of features for both PCA and PCA-based BB-BC technique as shown in Fig. 1 It is clear from the figure that recognition rate is directly proportional to the number of features. It is evident that PCA-based BB-BC algorithm outperforms single-handed PCA technique. It is worth noting that this hybrid approach quickly converges to maximum rate with less number of features. It is observed that the recognition rate saturates to its maximum from 6 features and onwards. It implies that less number of features targets to maximum recognition rate resulting in better performance.

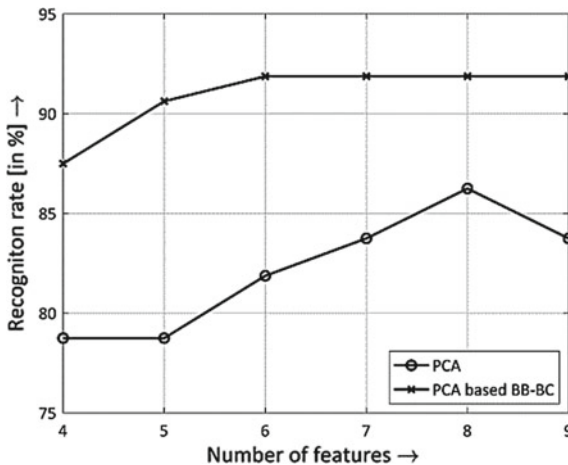


Fig. 1 Recognition rate depending on the number of selected features using a solely implemented PCA and hybrid PCA-based BB-BC algorithm

The improved performance of the hybrid approach is justified by the selection criteria of features. In the single-handed PCA technique, the features are selected in a consecutive manner. However, the hybrid approach follows distinct criteria for feature selection. For instance, when both techniques were allowed to choose four features, single-handed PCA techniques selected 1, 2, 3 and 4, and hybrid techniques picked 2, 4, 6 and 7. This process resulted in a higher recognition rate of the hybrid technique by approximately 9%.

4 Conclusion

In this work, hybrid approach, namely PCA-based BB-BC algorithm, is elaborated. It is applied on ORL face database. The results are motivating as it concludes that the application of BB-BC algorithm for feature selection gives best recognition rate compared to PCA method for the same number of features in subset. It also signifies that the image can be reconstructed by lesser number of features using this hybrid approach. It gives rise to the possibility that better results could be achieved solely using soft computing tools, which is yet to be investigated.

References

1. P.N. Topiwala, *Wavelet Image and Video Compression* (Kluwer Academic, Norwell, USA, 1998).
2. D. Anthony, E. Hines, J. Batham, D. Taylor, A comparison of image compression by neural network and principal component analysis. *IICNN Int. Joint Conf. Neural Netw.* **1**, 339–344 (1990)
3. Q. Du, J.E. Fowler, Hyperspectral image compression using jpeg2000 and principal component analysis. *IEEE Geosci. Remote Sens. Lett.* **4**(2) (2007).
4. P.R. Oliviera, R.F. Romero, L.G. Nonato, J. Mazucheli, Techniques for image compression: a comparative analysis, in *IEEE Sixth Brazilian Symposium on Neural Networks* (2000), pp. 249–254
5. D. Kumar, C.S. Rai, S. Kumar, Dimensionality reduction using SOM based technique for face recognition. *J. Mult* **5**(1) (2008)
6. D. Kumar, C.S. Rai, S. Kumar, Principal component analysis for data compression and face recognition. *INFOCOM J. Comput. Sci.* **7**, 48–59 (2008)
7. H.J. Holz, M.H. Loew, Relative feature importance: a classifier independent approach to feature selection, in *Pattern Recognition in practice IV* ed by E.S. Gelsema, N.L. Kanal (Elsevier, Amsterdam, 1994), pp. 473–487
8. N. Kwak, C.H. Choi, Input feature selection by mutual information based on parzen window. *IEEE Trans. Pattern Analys. Mach. Intell.* **24**(12), 1667–1671 (2002)
9. I.T. Jolliffe, *Principal Component Analysis* (Springer Verlag, 1986)
10. M. Turk, A. Pentland, Eigenfaces for Recognition. *J. Cogn. Neurosci.* **3**(1), 71–86 (1991)
11. K. Erol Osman, E. Eksin, New optimization method: big bang-big crunch. *Adv. Eng. Softw.* **37**, 106–111 (2006). (Elsevier)

OntoVidRec: A Staged Knowledge Aggregation Scheme for Annotations-Based Video Retrieval Using Ontology Matching



Gerard Deepak and A. Santhanavijayan

Abstract The amounts of multimedia data on the Web have increased in recent times and there are very few approaches for video recommendations. There is a need for a semantically driven video recommendation system as the current structure of the Web is progressing into a rational Semantic Web. The increase in the number of videos has also enhanced the annotations and the metadata associated with the video recommendation system. Due to the luxury of the availability of an increase in the amounts of metadata on the Web specifically for the video contents, an annotations-based video retrieval scheme by aggregating the meta information of the video is a mandated requirement as it can reduce the computational complexity in video recommendation than the traditional content-based video recommendation systems. In this paper, an OntoVidRec framework which aggregated video-related meta information from the query, user profiles, and the dataset has been proposed. The approach builds formal ontologies from the individual sources and encompasses a strategic model for Ontology Matching to yield the most appropriate Query relevant Ontological Entities that are semantically matched with the video annotations and are recommended based on the semantic similarity and the Kullback-Leibler divergence measure. An overall F-Measure of 95.37% has been achieved by the OntoVidRec which is the best in class performance for such systems.

Keywords Annotations-based video retrieval · Ontologies · Recommendation systems · Semantic similarity · Video mining

1 Introduction

Video Recommendation from the World Wide Web in the era of Semantic Web is required owing to the exponential increase in the multimedia contents on the Web. The Video contents tend to increase the complexity than any other multimedia content owing to their large frame size and incorporation of audio into them. Video Analysis

G. Deepak (✉) · A. Santhanavijayan
Department of Computer Science and Engineering, National Institute of Technology,
Tiruchirappalli, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_73

957

is computationally expensive, and the recommendation of videos based on the video descriptions and annotations can reduce the complexity of the recommendation and yield satisfactory results. Most of the existing video recommendation systems are either content-based video recommendation systems with a very high computational complexity or consider only the titles of the video or the annotations of the video or descriptions. The latter which are text-driven do not perform well and usually end up being polysemous or have a low coverage without addressing the issues of ambiguity. The Learning-Based Approaches which select text features alone do not tackle the problem of ambiguity and still have a low coverage rate. The Hybrid Learning Approaches which consider the textual and the content features also lag in their performance owing to the fact that the video contents are a collection of frames that are in constant motion and learning from them does not yield results that are correct or satisfactory. Due to all these reasons a text-driven video recommendation approach but that considers every possibility of the text sources as indicator or clue terms is the solution. The best in class solution is to build an ontology from various sources to capture the relevance of recommendation from different perspectives. A Query Sourced Ontology, a user information ontology, and the source data-driven ontology can be modeled and recommendations can be done using some semantic strategies which will transform the problem of Video Recommendation into a paradigm of Ontology-Driven Semantic Search for the Videos which ensures it as an inferential scheme over a learning-driven paradigm.

Motivation: Video Recommendation has emerged to be necessitated in the current times where the multimedia contents of the Web is exponentially expanding. With several video recommendation platforms competing against each other, a computationally less expensive and yet a full cover video recommendation framework which tackles the problems of polysemy, cold start, and ambiguity is the need. Due to the complexity involved in the content-based video recommendation techniques and their lag in performing effectively, an annotation-based video recommendation approach that takes into consideration, the textual contents from several sources and perspectives is required, such that the recommendations are non-ambiguous with a contextual semantic search and performs efficiently.

Contribution: A strategic approach for an annotations-based video retrieval, the OntoVidRec has been proposed. OntoVidRec is a full cover search that tackles the problem of ambiguity, polysemy, cold -start and serendipity in semantic search. The approach makes use of the Query and formulates an ontology by aggregating Lexical and Auxiliary Knowledge to yield a Query Sourced Ontology. Similarly, a User Profile-based Ontology and the Dataset-Derived Ontology is formalized. The approach encompasses the SemantoSim measure and the Kullback-Leibler Divergence for Ontology Matching of the Query Sourced Ontology with the other ontologies in a systematic manner to yield the most recommendable tags which are used for recommending the videos to the user. An F-Measure of 95.18% has been achieved with a coverage of 0.98 which makes it a full cover search.

Organization: This paper is organized as follows. Section 2 provides an overview of the related research work. Section 3 presents the Proposed System Architecture. Implementation is discussed in Sect. 4. Performance Evaluation and Results is discussed in Sect. 5. This paper is concluded in Sect. 6.

2 Related Work

Yashar et al. [1] have put forth a strategy for recommending videos focusing on the stylistic features from the video. This approach not only facilitates content-based extraction of videos from full-length videos, but the approach is capable to extract features from an abstractive form of videos like the trailer of the full video. Diaz et al. [2] have proposed a cognitive approach for video recommendations based on the emotional reactions of the viewer. Shang et al. [3] have devised an approach that recommends a micro-video based on inferences made from Big Data using parallel computing as a scheme. The parallel computing is achieved using Hadoop and MapReduce frameworks and based on the inferences made from the analysis of the big data, the videos are recommended. Checkley et al. [4] have devised a scheme where videos are recommended in a dynamic environment involving video sharing by extracting the keywords from the title of the video and compares with the words in the target video by computing its similarity score and ranking it. Gao et al. [5] have proposed a dynamic RNN model to capture the changing user interests over time and yield a personalized video recommendation framework. Liu et al. [6] have proposed the multi-info-based fusion model for video recommendation based on the integration of the user rating and the textual data of the video names the description and the genre. The approach uses the Jaccard Similarity Measure for comparing the similarity between the items. Zhou et al. [7] have improvised the performance of video recommendations by encompassing the usage of social interactions of the user. Cai et al. [8] have proposed a multi-view learning framework for the recommendation of videos that focuses on class label querying. The approach proposes an MVAL algorithm for cutting down the annotation cost, and also encompass informative scenarios into the approach. Wei et al. [9] have proposed a mixed model for a collaborative filtering-driven video recommendation system by making use of user information and similarity, K-means clustering, video genetic structure for yielding style and regional preferences. Yongxia et al. [10] have proposed a relative algorithm for recommending videos which integrates a user-based collaborative filtering, item-based collaborative filtering, and a model-based collaborative filtering for an efficient recommendation of videos. Nitin et al. [11] have proposed a mechanism of concept-based recommendation of videos by combining a technique comprising of ranked intersection filtering and a foreground-based concept co-occurrence matrix which is a content-based video recommendation system that uses deep CNNs. Liu et al. [12] have devised an approach that recommends videos based on tags that are enhanced using a graph-based neural network for ranking of tags. The approach also assimilates a neighbor similarity scheme for loss estimation which facilitates encoding the

varied user preferences into possible representations of nodes. In [13–18] Ontologies have played a vital role in semantic enrichment of tags or vocabularies and ontology focused mechanisms have incorporated intelligence into the approach to yield the best in class recommendations in various scenarios.

3 Proposed System Architecture

The architecture of the proposed Semantic Aware Video Recommendation Framework is depicted in Fig. 1. The proposed framework is a composition of Ontology Modeling from three main sources namely the Query Sourced Ontology, User-Driven Ontology, and the Dataset-Derived Ontology. The approach looks quite significant, but this is the first of its kind which matches a single source Ontologies with Ontologies from two different sources for facilitating semantic search. The Query Sourced Ontology is obtained from the User Query which is input into the system. The user query undergoes simple pre-processing namely the Tokenization, Lemmatization, and the stop word removal. The pre-processed user query is formulated as a query word set. The query word set is subject to Named Entity Recognition for predicting the context of the user query, and then the query word set is subject to synonymizing using the WordNet 3.1.

Further to overcome the polysemy problem and if the query word is polysemous, then its equivalent Hypernyms, Hyponyms, and Meronyms are loaded based on the context of the query word. Further, the synonymized query word with hypernym, hyponym, and meronym aggregation based on the context is linked to the Wikidata through the SPARQL Endpoint to aggregate real-world knowledge based on the query words which have been synonymized to yield a taxonomy of Query Words. The Query Word Taxonomy is further Reasoned using the Pellet Reasoner and is formalized into a Query Sourced Ontology.

The User Profile Information is elicited, analyzed to understand the user preferences, and uncover the details concerning the categories and the nature of video that the user wishes to watch. The Web Usage Data from the user profile and the channel subscription information from the user profile is extracted, pre-processed based on Tokenization, Lemmatization, and stop word removal. The frequent terms from the Web Usage Data and the subscription are extracted and further synonymized. The terms are prioritized based on the frequent terms and are further subject to synonymization. The SPARQL Endpoint for Wikidata is used for Knowledge Leveraging and Aggregation. Furthermore, the terms are reasoned out and formalized to yield a User-Driven Ontology. The dataset is first pre-processed based on the Title of the video, Video Description, and the Video Metadata to yield descriptor and indicator terms. Among all the contents in the dataset, the focus is mainly on the Video Title, Annotations of the video, and the Metadata description of the video content. The pre-processed dataset yields video descriptor and indicator terms which are passed onto the reasoner for correlations among the contents. Further to this the term descriptor and indicators are linked with the Wikidata API to leverage real-world knowledge.

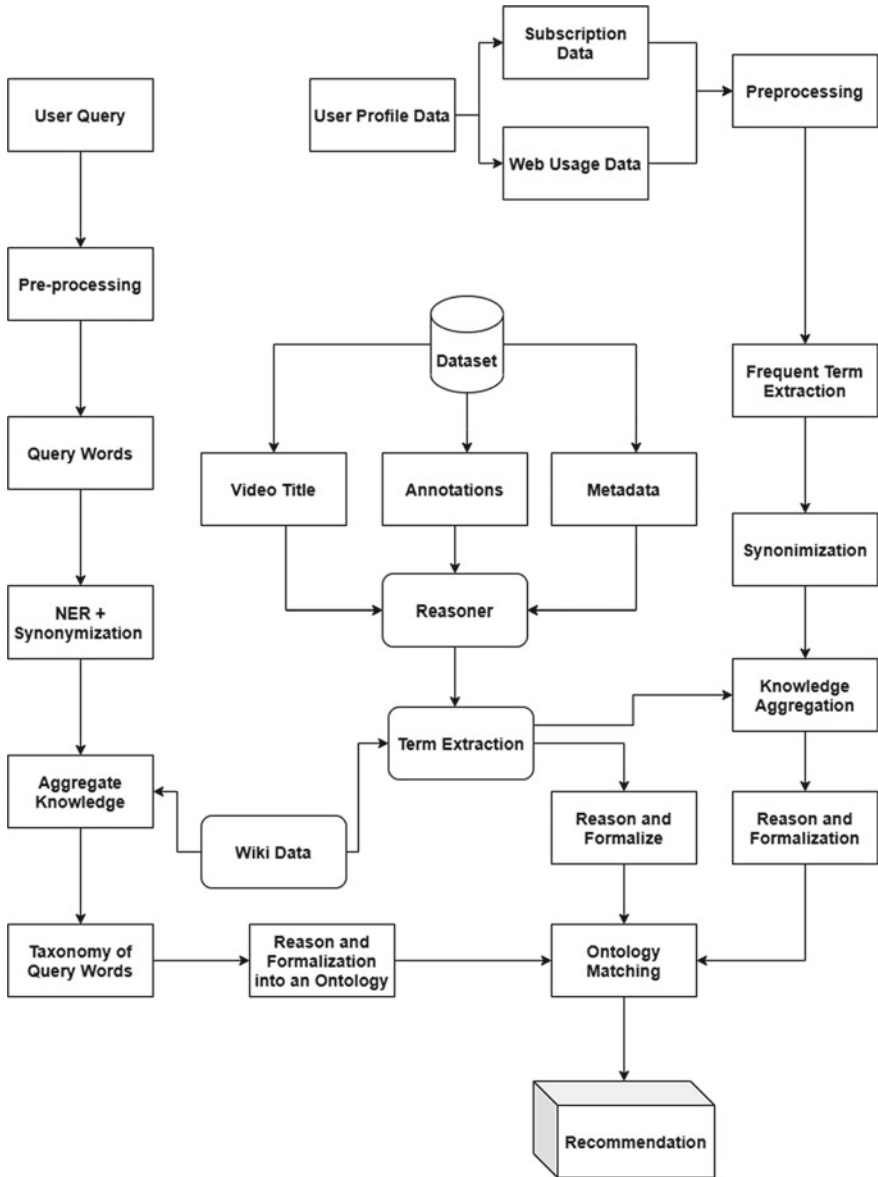


Fig. 1 Proposed system architecture

Further to this the terms are reasoned out and formalized into an ontology termed as the Dataset-Derived Ontology. The ontological terms are populated based on leveraging of the knowledge contents from the standard real-world knowledge base which is the Wikidata.

Once the Query Sourced Ontology, User-Driven Ontology, and the Dataset-Derived Ontology are formulated from the Query, User Profile Information, and the Dataset, Ontology Matching is performed by encompassing two constraints, namely the Kullback-Leibler (K-L) Divergence and the SemantoSim Similarity Measure. The Query Sourced Ontology is used as a source Ontology or the key for Matching. Ontology Matching is achieved in two stages namely where the SemantoSim measure is computed with a threshold of 0.5 among the concepts of the Query Sourced Ontology and the User-Driven Ontology; the matching concepts along with its two immediate neighbors with a direct link are alone retained in the ontology. Now the K-L Divergence [19] is computed between the Query Sourced Ontology and the new ontology with matching nodes and their neighbors with a 50% threshold. The Matching nodes from the User Profile Ontology are now Merged with the Query Sourced Ontology. Furthermore, based on the Merged Ontology as the node, a similar procedure is followed to retain the nodes in the Dataset Ontology and the retained nodes are matched with the newly Merged Query Sourced Ontology. The videos with data descriptors that are semantically similar to the newly Merged Query Sourced Ontology are recommended based on the increasing order of the semantic similarity.

The SemantoSim [20, 21] Measure is depicted in Eq. (1) which is proposed which depicts the semantics based on the probability of individual occurrence and the probability of co-occurrence of the terms in the web corpus or the text corpus. Equation (2) depicts the K-L Divergence as the information divergence which is a probabilistic distance model that represents the distance between a pair of probabilistic distributions [22].

$$\text{SemantoSim}(x, y) = \frac{\text{pmi}(x, y) + p(x, y) \log[p(x, y)]}{[p(x) \cdot p(y)] + \log[p(y, x)]} \quad (1)$$

$$K - L \text{ Divergence}(p||q) = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)} \quad (2)$$

4 Implementation

The experimentations were conducted on the YouTube-8M dataset with added descriptions of the videos based on the annotations in the Multilabel dataset. Apart from the annotations, textual descriptions based on the Categorizations in the dataset were used to add descriptions to the video using customized JAVA crawlers. To get user profile and subscription information of the user, 427 users were given the annotations, labels, and the video information in the dataset and were asked to browse,

binge watch videos based on their choices but restricted to their topics for 120 days in a separately created YouTube profile. Further, they were asked to subscribe to the channels they were interested in and that they thought belonged to the domain without deviance. At the end of 120 days, the user profile information which comprised of the Web Usage Data of the user and the channel subscription information were extracted and were used for the experimentations. The Terms were formulated into Taxonomy and then Ontological Structures using the OntoCollab [23, 24] framework as it supports dynamism and has the facilitation to link itself to the Wikidata API through an intrinsic SPARQL Agent. The formulations and steps in the proposed OntoVidRec algorithm are represented as Algorithm 1.

Algorithm 1: Proposed OntoVidRec Algorithm for Video Recommendation
<p>Input: The multi-word query Q which is input, User-Profile Information U_p, and the Dataset D</p> <p>Output: Query Centric, User-Relevant, and Data centric videos</p> <p>Begin</p> <p>Step 1: Tokenize and lemmatize the input query Q. Eliminate the stop words from Q to obtain a list of query words q_w.</p> <p>Step 2: The q_w is subject to NER and is synonymized using WordNet 3.1. The hypernym-hyponym-meronyms are identified for context determination for polysemous words to yield set SQ_w</p> <p>Step 3: for each term in SQ_w Load auxiliary knowledge from Wiki Data Return taxonomy of Query Words T_q</p> <p>Step 4 : Reason and Formalize T_q into a Query Sourced Ontology QSOnto</p> <p>Step 5: Load the Subscription Information and Web Usage Data from U_p, Pre-process and extract frequent terms and further Synonymize to yield set UP_s</p> <p>Step 6 : for each term in UP_s Load auxiliary knowledge from Wiki Data Return taxonomy of Query Words User Profile T_u. end for</p> <p>Step 7: Reason and Formalize T_u into a User Driven Ontology UDOnto</p> <p>Step 8: Extract the Video title, Annotations and Metadata descriptions from D, reason it using a pellet reasoner, Aggregate with Wiki Data to yield a data driven Taxonomy T_d</p> <p>Step 9: Reason and Formalized T_d into a DSOnto</p> <p>Step 10: Using SemantoSim and K-L Divergence perform Ontology Matching between the QSOnto and UDOnto strategically. Further match the resulting Ontology with the DSOnto and yield the matching videos to the tags and labels in the Dataset.</p> <p>End</p>

5 Results and Performance Evaluation

The YouTube-8 M dataset was used as a base dataset but it has been enriched based on the addition of descriptions and incorporation of user information to it, and the experimentations were conducted on the newly resultant dataset. The performance of the OntoVidRec was evaluated using Precision, Recall, Accuracy, F-Measure, and False Discovery Rate as the suitable metrics. The OntoVidRec has served as a full cover search as it takes into consideration the query details, user details and also is driven by the data in the dataset which makes it quite efficient and solves the serendipity and the polysemy problems that are prevalent in web search. The experimentations were conducted for 1748 queries which were given to the users who participated in the user profile information collection and top-10 relevant items for recommendation were recorded for ground truth collection. Each of the query results was correlated with the ground truth for relevance where the true positives and the false positives were estimated for each of the query results and the overall average performance was evaluated.

From Fig. 2 it can be inferred that the OntoVidRec yields an average precision of 93.89%, an average recall of 96.87%, an average accuracy of 94.33%, and an average F-Measure of 95.37%. From Fig. 3 it is evident that the proposed OntoVidRec yields a high Coverage of 0.98 and a low FDR of 0.07. The reason for high Precision, Recall, Accuracy, F-Measure, and Coverage is not just the fact that the proposed OntoVidRec is Tag driven or due to the reason that it is driven by text and not the video content but the main reason is that it is quite selective in integrating the text contents from several heterogeneous sources. Most importantly, the user query is transformed into an ontology which is staged by accumulating knowledge based on Synonymization, Hypernym–Hyponym extraction, Wikidata based Knowledge Aggregation to yield a Query Sourced Ontology. Furthermore, the Subscription Information and the Web Usage Data from the user profile information is formalized into a User-Driven

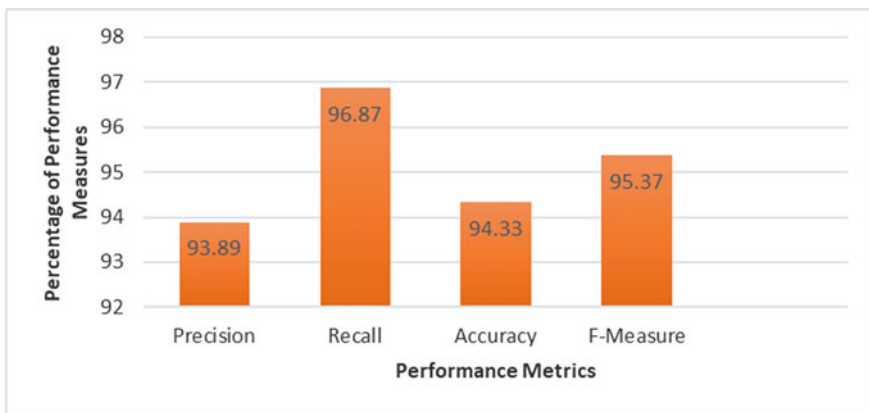


Fig. 2 Performance evaluation of the OntoVidRec

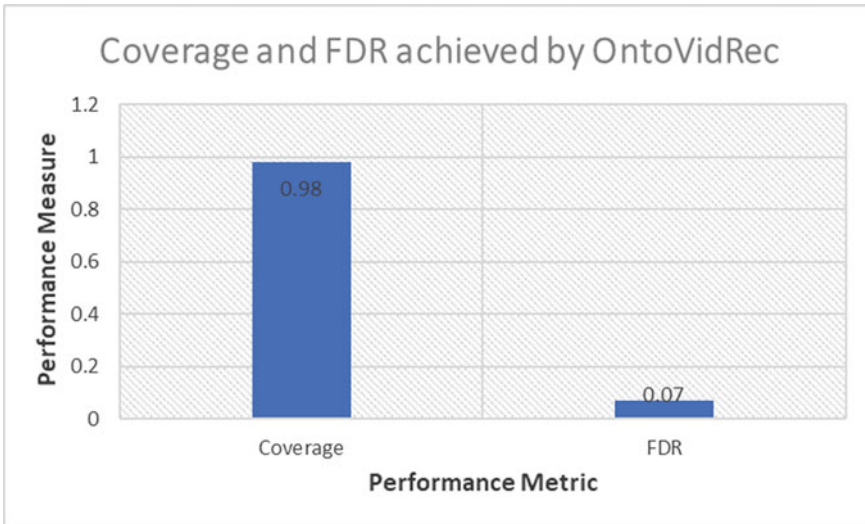


Fig. 3 Coverage and FDR of OntoVidRec

Ontology by encompassing frequent term extraction, Synonymization, and Wikidata based Knowledge Aggregation. Also, an ontology is formalized from the Dataset by pre-processing, synonymization, and knowledge aggregation. The recommendation is based on a strategic scheme of Ontology Matching based on the SemantoSim measure and the K-L Divergence as objective functions. Ontology Matching enables the semantics between Ontologies from varied perspectives and vantage points rather than a specific single source. This not only improves the context of recommendation but also enhances the coverage of the semantic search. The provision of dynamic auxiliary knowledge from the Wikidata Knowledge Source helps in overcoming cold start problem, ambiguity problem, context irrelevance in web search and also provides a deep sense of understanding of the query words and the probable entities that are relevant to the query. This enhances the Precision, Recall, Accuracy, and F-Measure scores and is also a reason for a low FDR value.

The proposed OntoVidRec Framework is baselined with the state of art video recommendation frameworks namely the RABCF [10] approach, SC-CNN [11] approach, and the GNN [12]. The baseline methods were evaluated in the exact same environment of the proposed OntoVidRec for the same dataset and the exact same number of queries. It is seen that from Table 1 that RABDC yields an average precision of 83.64%, an average recall of 80.61%, an average accuracy of 82.04%, F-Measure of 82.09%, and an FDR of 0.17. The SC-CNN has yielded an average precision of 88.78%, an average recall of 90.89%, an average accuracy of 89.21%, an average F-Measure of 89.81, and an FDR of 0.12. The GNN has furnished an average precision of 91.48%, an average recall of 93.77%, an average accuracy of 92.18%, and an average F-Measure of 92.61 with an FDR of 0.09. However, the performance of OntoVidRec is higher than the baseline strategies with an average

Table 1 Comparison of performance of hybrid semantic algorithm with other systems

Search technique	Average precision (%)	Average recall (%)	Accuracy (%)	F-measure	FDR
RABDC [10]	83.64	80.61	82.04	82.09	0.17
SC-CNN [11]	88.78	90.87	89.21	89.81	0.12
GNN [12]	91.48	93.77	92.18	92.61	0.09
OntoVidRec	93.89	96.87	94.33	95.37	0.07

Precision of 93.89%, an average Recall of 96.87%, an average accuracy of 94.33%, and an average F-Measure of 95.37% with the low FDR of 0.07.

The RABCF is a collaborative filtering-based approach that infuses the user-based collaborative filtering, item-based collaborative filtering, and a model-based collaborative filtering algorithm. However, there is no auxiliary knowledge fed into the system, but it is driven by the profile of the users' and the collective ratings of the items. The lack of structured knowledge makes it lag to a small extent. The SC-CNN integrates ranked intersection filtering with a foreground-based concept co-occurrence matrix which uses the deep CNNs. The SC-CNN model is a semantic concept-driven content-based model which employs the deep CNN which makes the approach computationally expensive and although semantics are imbibed into the methodology, the absence of background knowledge from standard knowledge bases is evident. The GNN is a tag-based scheme which uses a graph-based neural network for tag ranking and recommends the videos based on the tags ranked. Although this method is quite efficient, there is still a complexity in the usage of graphs with neural networks, and most importantly, the neighborhood similarity scheme makes it much more computationally expensive.

The proposed OntoVidRec framework suffices all the lacunae in the baseline systems and ensures that sufficient auxiliary knowledge is imbibed into the system in the form of Synonyms and background information from Wikidata Knowledge Base for the formalization of the Query Sourced Ontology, User-Driven Ontology, and the Dataset-Derived Ontology. Moreover, the Query Sourced Ontology has strategically matched with the User-Driven Ontology and the Dataset-Derived Ontology using SemantoSim measure and the K-L Divergence. Ontology Matching is computationally less expensive when compared to the graph-based neural network scheme and the content-based video recommendation which incorporates CNN. Moreover, the OntoVidRec framework amalgamates Query Information, User Profile Information consisting of the Subscription Information and the Web Usage Data, and an ontology from the data in the Dataset. Apart from this synonym from WordNet 3.1, Homonyms, Hypernyms, and background knowledge from Wikidata make OntoVidRec quite rich in its knowledge density which enhances the coverage and the performance in terms of Precision, Recall, Accuracy, and F-Measure.

6 Conclusions

A novel framework for an ontology focused video recommendation, the OntoVidRec has been proposed. The OntoVidRec serves as a multifaceted versatile video recommendation approach by formalizing ontologies from varied perspectives namely the Query Sourced Ontology, the ontology formulated from the user profiles, and the Dataset-Derived Ontology. The approach matches Ontologies based on the SemantoSim measure and the K-L Divergence to yield the best in class videos based on the video annotations or tags. The strategy involves NER, Synonymization, hypernym-hyponym identification, and Wikidata is used for knowledge aggregation which ensures that the approach is free from context irrelevance and ambiguity problems. The proposed OntoVidRec yields an average accuracy of 94.33% with a very low FDR of 0.06 which ensures that OntoVidRec is the best in class approach for annotations-based video recommendation.

References

1. Y. Deldjoo, M. Elahi, P. Cremonesi, F. Garzotto, P. Piazzolla, M. Quadrana, Content-based video recommendation system based on stylistic visual features. *J. on Data Semant.* **5**(2), 99–113 (2016)
2. Y. Diaz, C.O. Alm, I. Nwogu, R. Bailey, Towards an affective video recommendation system, in *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)* (IEEE, 2018), pp. 137–142
3. S. Shang, M. Shi, W. Shang, Z. Hong A micro-video recommendation system based on big data, in *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*. (IEEE, 2016), pp. 1–5
4. G. Checkley, L.L.C. Google, Video recommendation based on video titles. U.S. Patent 10,387,431 (2019)
5. J. Gao, T. Zhang, C. Xu, A unified personalized video recommendation via dynamic recurrent neural networks (2017)
6. Y. Liu, G. Zhang, X. Jin, H. Yuan Multi-info fusion based video recommendation system. *J. Phys. Conf. Ser.* **1229**(1), 012010 (2019) (IOP Publishing)
7. X. Zhou, L. Chen, Y. Zhang, D. Qin, L. Cao, G. Huang, C. Wang, Enhancing online video recommendation using social user interactions. *VLDB J.* **26**(5), 637–656 (2017)
8. J.J. Cai, J. Tang, Q.G. Chen, Y. Hu, X. Wang, S.J. Huang, Multi-view active learning for video recommendation. *IJCAI*, 2053–2059 (2019)
9. W. Zhang, Y. Wang, H. Chen, X. Wei, An efficient personalized video recommendation algorithm based on mixed mode, in *2019 IEEE International Conferences on Ubiquitous Computing & Communications (IUCC) and Data Science and Computational Intelligence (DSCI) and Smart Computing, Networking and Services (SmartCNS)* (IEEE, 2019), pp. 367–373
10. Y. Zhao, B. Zhang, Research on recommendation algorithms based on collaborative filtering. *J. Phys. Conf. Ser.* **1237**(2), 022094 (2019). IOP Publishing
11. N. Janwe, K. Bhoyar, Semantic concept-based video retrieval using convolutional neural network. *SN Applied Sciences* **2**(1), 80 (2020)
12. Q. Liu, R. Xie, L. Chen, S. Liu, K. Tu, P. Cui, B. Zhang, L. Lin, Graph neural network for tag ranking in tag-enhanced video recommendation, in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (2020), pp. 2613–2620

13. G. Deepak, B.N. Shwetha, C.N. Pushpa, J. Thriveni, K.R. Venugopal, A hybridized semantic trust-based framework for personalized web page recommendation. *Int. J. Comput. Appl.* **42**(8), 729–739 (2020)
14. G. Deepak, J.S. Priyadarshini, Personalized and Enhanced Hybridized Semantic Algorithm for web image retrieval incorporating ontology classification, strategic query expansion, and content-based analysis. *Comput. Electr. Eng.* **72**, 14–25 (2018)
15. G. Deepak, D. Kasaraneni, OntoCommerce: an ontology focused semantic framework for personalised product recommendation for user targeted e-commerce. *Int. J. Comput. Aided Eng. Technol.* **11**(4–5), 449–466 (2019)
16. C.N. Pushpa, G. Deepak, A. Kumar, J. Thriveni, K.R. Venugopal, OntoDisco: improving web service discovery by hybridization of ontology focused concept clustering and interface semantics, in *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)* (IEEE, 2020), pp. 1–5
17. A. Kumar, G. Deepak, A. Santhanavijayan, HeTOnto: a novel approach for conceptualization, modeling, visualization, and Formalization of domain centric ontologies for heat transfer, in *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)* (IEEE, 2020), pp. 1–6
18. S. Mandal, V.E. Balas, R.N. Shaw, A. Ghosh, Prediction analysis of idiopathic pulmonary fibrosis progression from OSIC dataset, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, (2–4 Oct. 2020), pp. 861–865. <https://doi.org/10.1109/GUCON48875.2020.9231239>
19. S. Haribabu, P.S. Sai Kumar, S. Padhy, G. Deepak, A. Santhanavijayan, N.D. Kumar, A novel approach for ontology focused inter- domain personalized search based on semantic set expansion, in *2019 Fifteenth International Conference on Information Processing (ICINPRO)* (Bengaluru, India, 2019), pp. 1–5, <https://doi.org/10.1109/ICInPro47689.2019.9092155>
20. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol. 661. ed. by M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh (Springer, Singapore, 2021) https://doi.org/10.1007/978-981-15-4692-1_58
21. S. Kullback, R.A. Leibler, On information and sufficiency. *Ann. Math. Stat.* **22**(1), 79–86 (1951). <https://doi.org/10.1214/aoms/1177729694.JSTOR2236703.MR0039968>
22. G.L. Giri, G. Deepak, S.H. Manjula, K.R. Venugopal, OntoYield: a semantic approach for context-based ontology recommendation based on structure preservation, in *Proceedings of International Conference on Computational Intelligence and Data Engineering* (Springer, Singapore, 2018), pp. 265–275
23. S. Mandal, S. Biswas, V.E. Balas, R.N. Shaw, A. Ghosh, Motion prediction for autonomous vehicles from lyft dataset using deep learning, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)* (30–31 Oct. 2020), pp. 768 – 773. <https://doi.org/10.1109/ICCCA49541.2020.9250790>
24. C.N. Pushpa, G. Deepak, J. Thriveni, K.R. Venugopal. Onto collab: strategic review oriented collaborative knowledge modeling using ontologies, in *2015 Seventh International Conference on Advanced Computing (ICoAC)*. (IEEE, 2015), pp. 1–7

An Approach for Retrieval of Text Documents by Hybridizing Structural Topic Modeling and Pointwise Mutual Information



K. Vishal, Gerard Deepak, and A. Santhanavijayan

Abstract As a result of the increase in the Internet usage, the data on the Web has been expanding exponentially. Since most of the data is unorganized and scattered globally, it is challenging to acquire the best-associated document to a user query. This article uses structural topic modeling along with pointwise mutual information and recurrent neural network to acquire the best-associated document to the query. In structural topic modeling, the idea used is that each word will have a probability to be appropriate to a topic. The topics belonging to each document are identified by using pointwise mutual information of the words that it comprises. It will be easy-going to discover the document of the topics, which user has asked for in his query. Here, recurrent neural networks are used to classify the query-related documents to yield better results. Also, pointwise mutual information has been applied for finding the similarity between words. This research used the RCV2 dataset for experimentation where the normalized discounted cumulative gain, accuracy, F-measure, and false detection rate were compared for measuring the performance of the model. The experiment's results show that the proposed model performs better than the baseline variations and the baseline models.

Keywords Document retrieval · PMI · RNN · Structural topic modeling

1 Introduction

Everyday there are millions of files being shared over the Web. Due to this, the documents and the textual data have been increasing, and furnishing the best results for a specific query from a huge pile of data, which is unorganized, is not an easy task. There are millions of users on the Internet worldwide who depend on the web pages to obtain the best document for their submitted queries. And this can only be achieved by filtering and classifying this ocean of data properly so that no incorrect

K. Vishal · G. Deepak (✉) · A. Santhanavijayan
Department of Computer Science and Engineering, National Institute of Technology
Tiruchirappalli, Tiruchirappalli, Tamil Nadu, India
e-mail: gerry.deepu@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
S. Mekhilef et al. (eds.), *Innovations in Electrical and Electronic Engineering*,
Lecture Notes in Electrical Engineering 756,
https://doi.org/10.1007/978-981-16-0749-3_74

969

results for any query is shown in the results. This kind of work is not possible to accomplish manually because of the humongous size of available data and due to its unorganized and unstructured nature. An effective and efficient way is needed to organize the data and extract the pattern from the existing data. Among the several available solutions, data mining and text mining (TM) are efficient and can be scaled for widespread implementation. This hybrid is an effective approach for working with data that contains text documents. Text mining includes techniques such as information retrieval, machine learning, classification, and many more.

Motivation: Semantic gap is the difference between the expected results by the user for a given query and the discovered results. Traditional methods for text-based search, which uses the relevance information to search for the documents, have a substantial semantic gap. These types of algorithms and approaches take time and for all search operations of specific queries and retrieving the related documents.

Contributions: In this paper, structural topic modeling with the pointwise mutual information framework is used to sort the documents by using the semantic gap with each of the search queries. Based on the words, a document contains topics that are determined which are associated with it. Data containing the distribution of topics and words and other data containing the distribution of documents and topics are created. These data are provided to the structural topic modeling (STM) [1]. Using the pointwise mutual information, similarity between a pair of terms is estimated to compute the similarity between a pair of terms.

Organization: This remainder of the paper is organized as follows. Section 2 provides an overview of the related work. Section 3 presents the proposed system architecture. Implementation is discussed in Sect. 4. Performance evaluation and results are discussed in Sect. 5. This paper is concluded in Sect. 6.

2 Related Work

Ensan and Bagheri [2] proposed a way for document retrieval by using the degree of relatedness of a document to the query. Kratzwald and Feuerriegel [3] proposed the way to give results to query after classifying the documents by neural networks that have been acquired by document retrieval method. The classification part will control the part of controlling the number of documents that are needed to be retrieved. To automatically organize the documents, Jacobi et al. [4] used latent Dirichlet allocation (LDA) by measuring the pattern of word occurrence. They used LDA topic modeling to analyze the news content and a large number of journalism texts. Ostrowski [5] used topic modeling in their paper to identify and classify the noteworthy topics from the Twitter messages. In his research experiments, these methods were found to be effective in the identification of sub-topics. They were also doing the classification of the large-size datasets. Onan et al. [6] used LDA, which is a probabilistic model to represent collections using term-based modeling, for representing the document collection. They used models as naïve Bayes, SVM logistic regression, and K-nearest

neighbor for the empirical analysis. Lane and Buglak [7] proposed a way for estimating and visualizing the topics in a collection of documents where each of the documents has more than one topic associated with it. This confines the periods for topics. In [8–20], several facets of semantic retrieval models have been discussed.

3 Proposed System Architecture

The proposed framework retrieves the related document, as shown in Fig. 1. Initially, query input from the user is tokenized and lemmatized. Before further processing, there is a need to perform the normalization of the text. Further, the synonyms are discovered using the WordNet API. The synonyms are found for the purpose of collecting more relatable documents. Alongside the Web usage information of the user is collected for getting the more terms user might be searching for. Further, these terms are fed to STM framework, which are being used for topic modeling. With these data, the STM framework, the social Web, and the Twitter API are also fed with Wikidata [21–23].

Structural Topic Modeling: Starting with a small introduction about the model that is proposed for the work, the name itself says that it is a topic model; like other conventional models, STM also uses the word count in the document to predict the topic to which it belongs to. In STM, the concept that each word having a probability to belong to a topic is used. Topic modeling methods often consider that each document is a bundle of words; each topic has a distribution of words; and

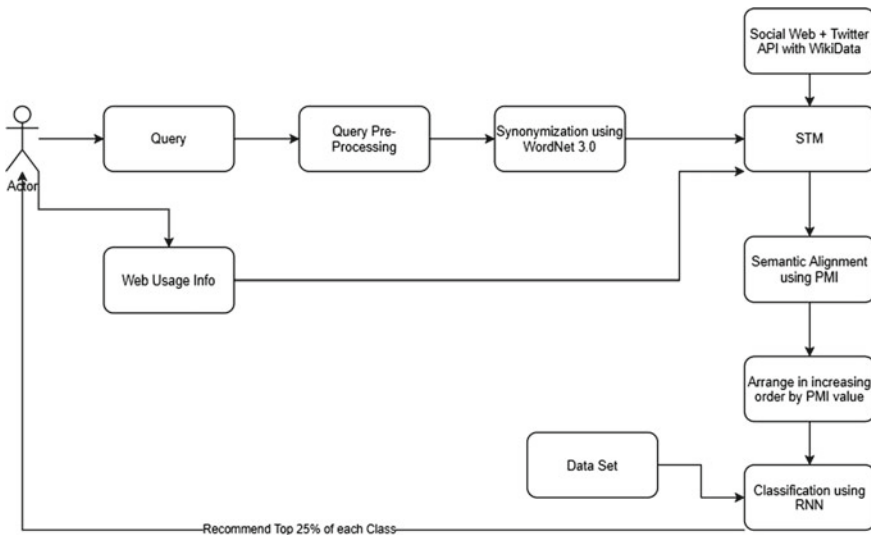


Fig. 1 Architecture of the proposed framework

hence, it can be inferred that each document is a combination of the topics. Herewith, it can be concluded that the documents mixture of one or more than one topic means each document comprises of several topics. This implies that the sum of all the proportion distribution of all the topics for a document will be one. And similarly, for a word probability distribution for a topic, the sum of the all the probabilities will be one.

Based on the words present, a document contains topics that are determined which are associated with it. Data containing the distribution of topics, words and other data containing a distribution of documents and topics are created. These data are fed into the STM framework. STM uses the concept that defines the extent by how much a topic is related to a document and words contained in it. This will be a function of the metadata associated with document data. Hence, it is obvious that topical prevalence covariate to the states of the extent of relatedness and topical contents. The proposed model amalgamates STM with pointwise mutual information to find the similarity between the word for topic identification. By using PMI, semantic alignment is achieved. Semantic alignment is shifting the data such that there is no bias for any data whatsoever. All the data has been made to refer to a single thing, which will reduce the biases of the procedure to specific data. Another reason for doing this is that many data will be related to different objects and phenomena, so the data to a single object or phenomenon can be combined. There is some more information about pointwise mutual information, which is used for semantic alignment.

Pointwise Mutual Information: The pointwise mutual information (PMI) uses the concept of probability to identify the extent to which two words are similar. The PMI between two words A and B is the likelihood of the presence of word A when it is known that word B is present in the document. It is a metric to measure the likelihood of two words being present in a document. The data returned by STM is rearranged using the PMI values and arranged in increasing order. This data, alongside the dataset, is fed to recurrent neural network for classification based on the user's query.

Recurrent Neural Network: Recurrent neural network (RNN) is a kind of neural network where the input to the present stage is given by the previous stage. In conventional neural networks, the input and input to different nodes of the network are unrelated. But the problem arises when a need arises to predict the next word for a sentence. To accomplish that, the earlier word that was present in the sentence must be known. This is where RNN comes into play where each word output is taken as input to the same node to predict the next word, i.e., the hidden state keeps some information about the sequence. RNN has a memory unit that keeps the record of information that has been collected and calculated. RNN uses the same parameters to input, performs the task on inputs, and generates output in the hidden layer. This reduces the complexity of the neural network. After classification by RNN, the top 25% of each class is delivered to the user for his query.

4 Implementation

The famous dataset Reuters Corpus Volume I has been used for experimenting with the proposed framework in the research. This data is widely used for text mining research. It contains data in XML format. Metadata is avoided so that it is impartial. Reuters Corpus Volume II (RCV2) has data of about 800,000 categorized news stories. This data is made by Reuters Ltd. for research purposes. In pre-processing, the XML tags are removed along with URL and metadata information. An offline dataset is also used for the experiment. Key terms are collected from the dataset by removing stop word and performing stemming. The algorithm shown in Table 1 has

Table 1 Proposed hybrid algorithm for document retrieval

```

Input: Query
Output: Related Documents from the dataset.
Begin
    Take query q as input
    #Preprocessing of query:
    Import General_Stop_words, Numbers, Punctuations
    #Tokenization:
    Tokenized_q = Splitting the terms in the query q
    #Normalization:
    For each term in Tokenized_q
        If term is in General_Stop_words or Numbers:
            Normalized_temp = Tokenized_q - term
        If term start or end with Punctuations:
            Normalized_q = Noramlized_temp Remove Punctuations
    #Synonymization:
    For each term in Normalized_q:
        Search for synonyms of term using WordNet 3.0
        Synonymized_q = Normalized_q + synonyms
    User_query = Synonymized_q
    #Structural Topic Modeling
        Collect Web_usage_info from user
        User_query + Twitter_API_wikidata as docs_data fed to STM model:
    For term in User_query:
        If term is in docs_data:
            Calculate probability of term in that doc
            Probability_matrix = probability of term in the document
        #Semantic Alignment using PMI:
        For terms in user_query:
            Find PMI for two terms
            PMI_values = PMI similarity
            Rearrange the PMI_values for increasing order
        #Classification using RNN:
    Take Dataset and PMI_values and Classify the data based on PMI values.
    Return Top 25% for each class
End
    
```

been employed for the research. The implementation was carried out using a system with 4 GB DDR3 RAM and Intel Core i3 fifth generation processor.

The proposed algorithm 1 takes the input of the user and pre-processes it to gain the key insights of the user asked query. To find the key points from the query, the general stop words and numbers were extracted and removed them from the query, which have been already tokenized, i.e., split terms from the query. After tokenization, the punctuations were also removed from the query. Following removal of stop words, the synonyms of the key terms of the query are found which have gone through the previous process. The synonyms are added to the user query as well. This data, alongside web usage information of the user and Wikidata, which have associated metadata to it, is provided to the STM model, which found the probability of the individual terms in the documents. Using the PMI, the similarity between the two terms is computed. This is used for semantic alignment. Finally, recurrent neural network is used for the classification of the dataset and returned the related documents.

5 Results and Performance Evaluation

The performance for the proposed approach was evaluated using the precision, recall, accuracy, F-measure, normalized discounted cumulative gain (NDCG), and the false discovery rate (FDR). From Table 2, it is evident that the proposed model performs better than the baseline models and the variations of the other models. The proposed model has 15.66% higher precision than that of baseline variation of LDA with Cosine 13.29% higher precision than the DRDLC. It also has 15.65% higher recall, 15.66% higher F-measure, and 15.65% higher accuracy than baseline variation LDA + Cosine, whereas 14% higher recall, 14.19% higher F-measure, and 14.19% higher accuracy than baseline model DRDLC. It is because topic models make use of more

Table 2 Comparison of the performance measures for the baseline and other variations

Model	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)
LDA + Cosine	73.12	76.14	74.60	74.63
LSI + Jaccard	76.81	79.32	78.04	78.06
LSI + PMI	77.17	79.87	78.50	78.52
LDA + PMI	75.14	77.89	76.49	76.52
STM + Jaccard	80.12	82.56	81.32	81.34
STM + PMI	81.89	83.17	82.53	82.53
STM + RNN	84.78	87.14	85.94	85.96
DRDLC	74.86	77.32	76.07	76.09
Proposed approach	88.78	91.79	90.26	90.29

sophisticated Bayesian techniques to find the probability that a document is associated with a given theme or topic. This means that the probabilities become increasingly accurate as more data are processed. Combined with PMI and RNN, it produces decent results.

From Fig. 2, it can be seen that the performance of the proposed approach outperforms the other models. The proposed model has about 0.17 higher NDCG than the baseline variation LDA + Cosine, 0.15 higher NDCG than LSI + Jaccard, 0.13 higher gain than LSI + PMI, whereas 0.07 higher NDCG than of baseline model of DRDLC, 0.08 more gain than STM + PMI, and 0.04 more gain than STM and RNN. The reason for the proposed approach to yield better results is that topic modeling offers several ways to summarize, arrange, and identify huge collections of information that consist of text. Also, in the proposed model, synonymizing was performed, which enabled to produce efficient results as the semantic gap is reduced and all the related vectors point to the same word.

From Fig. 3, it is evident that the FDR for the proposed model is 0.04 lower than the baseline variation, 0.02 lower than LSI + Jaccard, 0.01 lower than LSI + PMI,

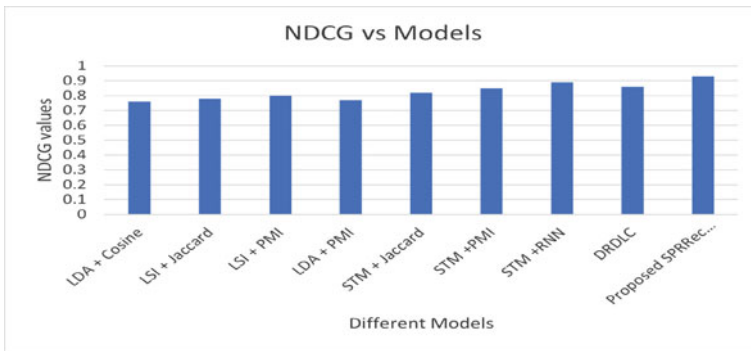
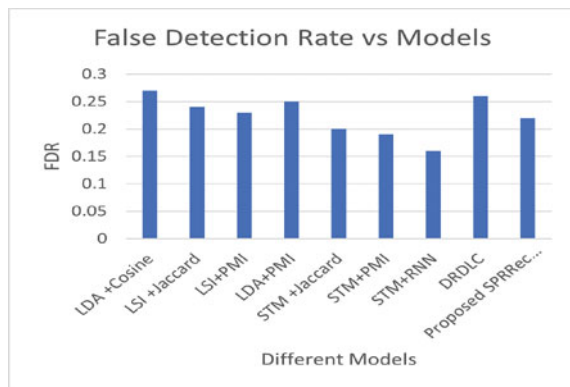


Fig. 2 Comparing the NDCG values for model variations and baselines

Fig. 3 FDR for model variations and baselines



and 0.03 lower than that of LDA + PMI, whereas 0.04 lower FDR than the baseline model DRDLC, 0.03 higher FDR than STM + PMI, and 0.06 FDR higher than the STM + RNN model. It is because topic modeling helps in discovering hidden topical patterns that are present in the collection of data interpreting documents according to these topics. And RNN keeps information during each cycle. Because of RNN's property of keeping records of previous inputs and outputs, it is widely used for time series analysis.

6 Conclusions

The proposed framework can find related documents efficiently and effectively. Unlike the traditional ways of searching for documents, this framework uses terms in the documents and queries to find the related documents. The results show that the proposed model has performed better than that of the baseline variation models and baseline models. Topic modeling, pointwise mutual information, and recurrent neural network have helped the model perform to its best in document retrieval. An overall F-measure of 90.29% has been achieved for the proposed hybrid methodology for document retrieval.

References

1. M.E. Roberts, B.M. Stewart, D. Tingley, Stm: an R package for structural topic models. *J. Stat. Softw.* **91**(1), 1–40 (2019)
2. F. Ensan, E. Bagheri, Document retrieval model through semantic linking, in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining* (2017, February), pp. 181–190
3. B. Kratzwald, S. Feuerriegel, Adaptive document retrieval for deep question answering (2018). ArXiv preprint [arXiv:1808.06528](https://arxiv.org/abs/1808.06528)
4. C. Jacobi, W. Van Atteveldt, K. Welbers, Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism* **4**(1), 89–106 (2016)
5. D. Ostrowski, Using latent dirichlet allocation for topic modelling in twitter, in *2015 IEEE International Conference on Semantic Computing (ICSC)* (Anaheim, CA, USA, 2015), pp. 493–497
6. A. Onan, S. Korukoglu, H. Bulut, LDA-based topic modelling in text sentiment classification: an empirical analysis. *Int. J. Comput. Linguistics Appl.* **7**(1), 101–119 (2016)
7. A. Lane, R. Buglak, BAE Systems Australia Ltd. System and method for identifying and visualising topics and themes in collections of documents. U.S. Patent Application 14/387, 268 (2015)
8. G. Deepak, A. Santhanavijayan, OntoBestFit: a best-fit occurrence estimation strategy for RDF driven faceted semantic search. *Comput. Commun.* **160**, 284–298 (2020)
9. C.N. Pushpa, G. Deepak, A. Kumar, J. Thriveni, K.R. Venugopal, OntoDisco: improving web service discovery by hybridization of ontology focused concept clustering and interface semantics, in *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)* (IEEE, 2020), pp. 1–5

10. A. Kumar, G. Deepak, A. Santhanavijayan, HeTOnto: A novel approach for conceptualization, modeling, visualization, and formalization of domain centric ontologies for heat transfer, in *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)* (IEEE 2020), pp. 1–6
11. G. Deepak, D. Kasaraneni, OntoCommerce: an ontology focused semantic framework for personalised product recommendation for user targeted e-commerce. *Int. J. Comput. Aided Eng. Technol.* **11**(4–5), 449–466 (2019)
12. Z. Gulzar, A. Anny Leema, G. Deepak, Pcrs: personalized course recommender system based on hybrid approach. *Proced. Comput. Sci.* **125**, 518–524 (2018)
13. S. Mandal, S. Biswas, V.E. Balas, R.N. Shaw, A. Ghosh, Motion prediction for autonomous vehicles from lyft dataset using deep learning, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)* (30–31 Oct. 2020), pp. 768–773. <https://doi.org/10.1109/ICCCA49541.2020.9250790>
14. G. Deepak, V. Teja, A. Santhanavijayan, A novel firefly driven scheme for resume parsing and matching based on entity linking paradigm. *J. Discrete Math. Sci. Crypt.* **23**(1), 157–165 (2020)
15. S. Haribabu, P.S. Sai Kumar, S. Padhy, G. Deepak, A. Santhanavijayan, N.D. Kumar, A novel approach for ontology focused inter-domain personalized search based on semantic set expansion, in *2019 Fifteenth International Conference on Information Processing (ICINPRO)*, (Bengaluru, India, 2019), pp. 1–5. <https://doi.org/10.1109/ICInPro47689.2019.9092155>.
16. G. Deepak, G.V.S.N. Naresh Kumar, S.Y. Bharadwaj, A. Santhanavijayan, OntoQuest: an ontological strategy for automatic question generation for e-assessment using static and dynamic knowledge, in *2019 Fifteenth International Conference on Information Processing (ICINPRO)* (IEEE, 2019), pp. 1–6
17. A. Santhanavijayan, D. Naresh Kumar, G. Deepak, A semantic-aware strategy for automatic speech recognition incorporating deep learning models, in *Intelligent System Design* (Springer, Singapore), pp. 247–254
18. M. Kumar, V.M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive Data Analysis for Energy Management of a Smart Factory Leading to Sustainability, in *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol. 661. ed by M. Favorskaya, S. Mekhilef R. Pandey, N. Singh (Springer, Singapore, 2021) https://doi.org/10.1007/978-981-15-4692-1_58
19. G. Deepak, et al., Design and evaluation of conceptual ontologies for electrochemistry as a domain, in *2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)* (IEEE, 2019)
20. G. Deepak, J.S. Priyadarshini, Personalized and enhanced hybridized semantic algorithm for web image retrieval incorporating ontology classification, strategic query expansion, and content-based analysis. *Comput. Electr. Eng.* **72**, 14–25 (2018)
21. G. Deepak, J.S. Priyadarshini, M.H. Babu, A differential semantic algorithm for query relevant web page recommendation, in *2016 IEEE International Conference on Advances in Computer Applications (ICACA)* (IEEE, 2016), pp. 44–49
22. I.S. Kaushik, G. Deepak, A. Santhanavijayan, QuantQueryEXP: a novel strategic approach for query expansion based on quantum computing principles. *J. Discrete Math. Sci. Crypt.* **23**(2), 573–584 (2020)
23. A. Santhanavijayan, D.N. Kumar, G. Deepak, A novel hybridized strategy for machine translation of Indian languages. *Soft Comput. Sig. Proc.* 363

Deep Belief Convolutional Neural Network with Artificial Image Creation by GANs Based Diagnosis of Pneumonia in Radiological Samples of the Pectoralis Major



Tathagat Banerjee, Dhruv Batta, Aditya Jain, S. Karthikeyan, Himanshu Mehndiratta, and K. Hari Kishan

Abstract In this paper, we delineate a comparative classification of Pneumonia using Machine Learning and Deep Learning models. The data used was the dataset of Chest X-Ray Images for Classification made available by Kermay (2018) with a total of 5863 images, with 2 classes: normal and pneumonia. For the purpose of correcting the class imbalance between normal and pneumonia images, we use General Adversarial Networks to generate pneumonia ridden images. The comparative classification is built on a final dataset of 19,784 images. Logistic regression, SVC, KNN, Random Forest and other machine learning models such as XgBoost and CatBoost are compared with deep learning models such as Mobile Net and VGG-16. The machine learning model is based on blob segmentation and detecting the difference between blobs of pneumonia ridden and normal individuals. Finally, a new deep learning model with convolutional and artificial neural networks has been proposed for the classification purpose which increases the accuracy significantly and has a classification report which is best suitable for medical analysis.

Keywords Neural networks · Machine learning · ANN · CNN · GAN · F-CNN · Transfer learning · Artificial intelligence · Computational power

1 Introduction

Pneumonia is the single leading cause of mortality in children and is a major cause of child mortality in every region of the world. It is a form of acute respiratory tract infection (ARTI) that affects the lungs. When an individual has pneumonia, the alveoli in the lungs are filled with pus and fluid, which makes breathing painful and

T. Banerjee (✉) · D. Batta · A. Jain · H. Mehndiratta
VIT-AP University, Amaravati, India
e-mail: banerjee.tathagat@vitap.ac.in

S. Karthikeyan · K. Hari Kishan
Faculty Computer Science Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India

limits oxygen intake; this particular stance can be useful in detecting the difference in blobs of a normal person and pneumonia ridden individuals.

The X-Rays and Chest Tomography using CT Scans are cross-sectional images of the body allowing the internal organs to be picturized. The images are further used to identify abnormalities present in a patient.

Plain X-Rays while generating enough information CT Scans provide an edge [1] over since they help detect various blobs in the images. Images of CT scan and plain X rays have been used from the labeled dataset provided. Images have two classes:-

1. Pneumonia Ridden Patients
2. Normal Healthy Patients.

The deep learning model which has been proposed in this paper does the analysis of the CT-Scan with an accuracy of around 98% with real-world data. The model is also not prone to overfitting. This model aims at correctly predicting pneumonia in an image of X-Ray of the chest of the test subject.

2 Literary Review

In Khobragade et al. [2] proposed that in the detection of pulmonary diseases [3] using chest radiographs was necessary to identify the problem present in the tract. The lung diseases are a major problem to human health and thus for this identification, the referenced paper proposes pulmonary segmentation, extraction of characteristics and classifying it using an Artificial Neural Network for the recognition of the disease. A simple image processing with an intensity-based method was used in conjunction with a method to detect the pulmonary limits and the discontinuities in it. Artificial Neural Networks were used for feed-forward [4] and backpropagation. The problem with this approach is that it uses chest radiographs and the highest infant mortality rate in developing countries due to pneumonia and in these countries, there is little infrastructure and doctors in rural areas to provide the necessary diagnosis.

Many researchers have believed in finding specific patterns for various lung diseases through ultrasound images of the respiratory tract. In Barrientos et al. [5] the paper presented a method for automatic diagnostics of pneumonia using ultrasound imaging of the thoracic cavity. The approach presented in the paper is based on the analysis of patterns present in 'rectangular segments' from the sonographic digital images. Utilizing pattern recognition, specific features from the characteristic vectors are obtained and then classified with standard neural networks. A training and testing set of positive and negative vectors were compiled.

The approach is problematic here since they have extracted the image of a single patient to either test or train and obtained a specificity of 100% which points toward a pre-processing conceptual error.

The paper [6] presented a way for implementing the Structural Co-occurrence Matrix approach to classify nodules as malignant nodules or benign nodules and also the level of malignancy. The structural co-occurrence matrix technique was applied

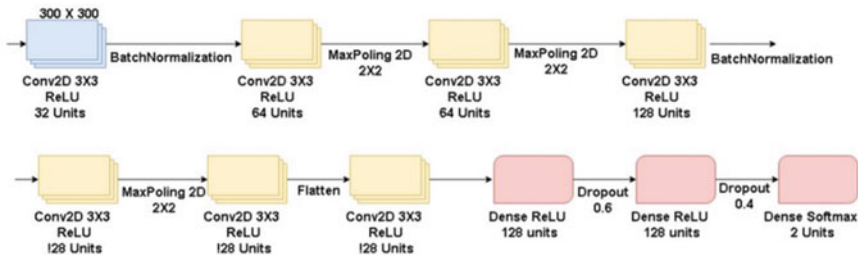


Fig. 1 Pipeline

to extract characteristics of the nodule images and classify them. The SCM [7] was applied in greyscale and images of the Hounsfield unit with four filters, creating eight different configurations. The classification stage used classifiers known as the multilayer perceptron, support vector machine, and k-Nearest Neighbors algorithm which were applied to two tasks: (i) to classify the nodule images as malignant or benign, (ii) to classify the nodules pulmonary lesions at the level of malignancy (1–5).

The approach while had a result of 96.7% for precision and F-score measurements in the first task and 74.5% accuracy and 53.2% F-score in the second task but did not achieve the much-required diagnosis of pneumonia pathological change identification and also lapses in terms of accuracy for determining benign and malignant nodules in the respiratory tract.

In Saraiva et al. [8]. A comparative classification of Pneumonia using Convolutional Neural Networks [9] is proposed in this paper and is an efficient method than the one proposed by Daniel S. Kermany in “Identifying Medical. Diagnoses and Treatable Diseases by Image-Based Deep Learning” [6].

The following is their architecture of their model which consists of the inherent problem during the general pre-processing phase as well as lacks the statistical approach in favor of deploying a larger network (Fig. 1).

The images in the training directory are more of the label pneumonia compared to that of normal. Therefore, class.

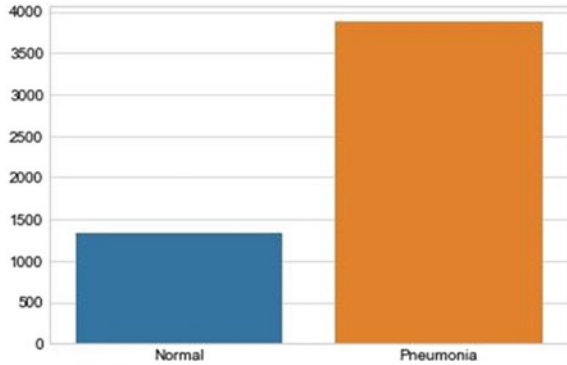
Balance is not maintained. Hence their accuracy of 95% is on a dataset with imbalanced classes Accuracy Matrix of this approach is displayed in the following Fig. 2.

3 Hardware and Software Requirements

The Complete experiment has been carried out on an 8 GB, Intel(R) Core (™) i3-6006U CPU @ 2.00, 1.99 GHz.

The operating system is 64-bit Operating System(×64-based processor), Windows 10 Home Single Language. Python 3.6.1 is used, with Tensorflow 2.0

Fig. 2 Distribution of classes



as software tools for the experimentation. Keras and Scikit Learn are used as dependency libraries for training and getting classification reports. Seaborn and Graphviz is used for plotting and obtaining schematics for the models. OpenCV 2.0 is used to read images and *tdqm* is used for the better graphical user interface.

About 40–48 h of total training time is given to the dataset to achieve good and industrial application level results.

The image quality of the data is set to as $128 \times 128 \times 3$ so that even when unsatisfactory and blurred image data is observed; the quality of results is maintained.

4 Transfer Learning Methodology

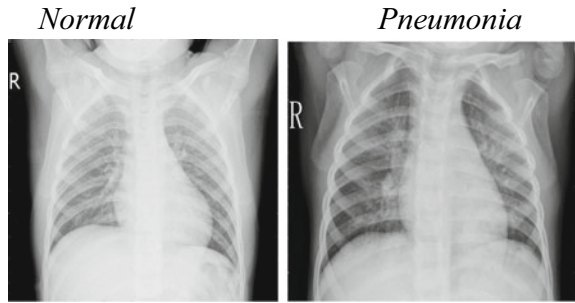
Initially, we trained the model with VGG [10], Mobile net, and Inception but none of the models gave us expected results when applied on real-time data. The highest accuracy was achieved by the Inception model which was 76% on training data but when applied on real-time data (test set) it could only attain an accuracy of 52.3%. We have modified the architecture as per our requirement by stacking up a few more layers of deep neural networks which are of type deep belief neural nets, we have named the architecture as Phoenix.

5 Data

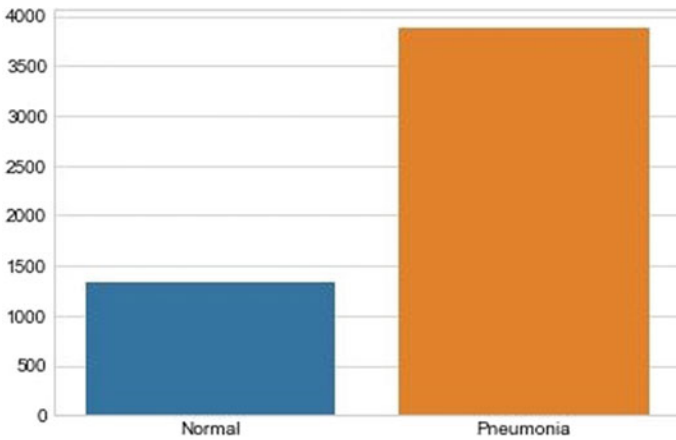
This dataset contains thousands of validated OCT and Chest X-Ray images.

All images in the dataset (Kermamy 2018) underwent treatment in order to remove all low-quality scans, as well as being classified by two specialist physicians and by a third-party specialist, in order to prevent any misclassification. As seen in other research on the same dataset, the training dataset lacks images for the ‘Normal’

Fig. 3 Sample images



label. Therefore, this creates a class imbalance and needs to be solved during the pre-processing. (Fig of Number of 'Normal' versus 'Pneumonia' Labelled images.



In the diagnosis of pneumonia, the alveoli become clogged with secretion and appear as a white spot on the chest radiograph. *Pulmonaryconsolidation* means that the alveoli in the lungs are filled with inflammatory fluid. In radiography, pulmonary consolidation dovetails to an opacity visible in the tract's radiography, that is, the whitish area. Which can be distinguished in Fig. 3 provided ahead.

6 Implementation

6.1 Machine Learning and Image Analysis

Machine learning [11] has been an important advancement in the field of predictive analysis over the last decade. However, in today's dilemma of predictive analysis both visible (Machine learning) and invisible analysis (Deep Learning) have become part

Fig. 4 Blob sampling



of continuous and prolonged debate. Thus, it is only wise to analyze the given data with both machine learning and deep learning. Our paper ideates the architectural stability and viability of the deep learning and neural network.

Blob segmentation is an important measure that enhances machine learning algorithms to learn. It is important to understand the relics and causes. Basically, blob detection means segmenting the portion of the image which is affected. Using Blob Detection would also create a simpler version of the algorithm to be utilized, whose results are not only generic in nature but simpler and faster to execute as well.

Blob is a group of connected pixels in an image that shares some common property. OpenCV provides a convenient way to detect blobs and filter them based on various characteristics. The algorithm is controlled by using parameters such as Thresholding, Grouping, Merging, and Center and radius calculation. In chest tomography and X-Rays of the thoracic cavity, we use parameters set to the specific filter of threshold and inertia. These filters deal with the possibility of the presence of a blob which may be the point of difference between normal and pneumonia images. The thresholding filter is based on the intensity of the blob present while inertia is based on the presence of elongated shape existing in the image which may be a factor. In the following figure, the two images show the blob difference of a healthy person and one suffering from pneumonia (Fig. 4).

The main principle mathematical concept of Blob detector is Laplacian of Gaussian. Given an input image $f(x, y)$, this image is convolve d by a Gaussian kernel.

$$g(x, y, t) = \frac{1}{2\pi t} e^{-\frac{x^2+y^2}{2t}}$$

After applying, the laplacian function we get the following equation:-

$$\nabla^2 L = L_{xx} + L_{yy}$$

This provides us with the filter through which our algorithm detects the blobs and the blob segmentation is also pursued through the same mathematical basis.

Even before any pre-processing measure, it is important to build strong and big image data, for which we used deep learning measure of unsupervised learning which is GAN (Generative Adversarial Networks) [12–14] and VAE (VariationAuto encoders) to prepare artificially equipped and balanced images so, that the medical stability of more ‘pneumonia’ images does not affect the working ideology and potential of any machine learning algorithm. This way, we not only up-sample the data but also do not create redundancy by using image generation techniques. The pre-processing phase involves reading the image file names with the help of OpenCV, then labeling all the files in a data frame along with a target variable—Normal, Pneumonia. Finally, shuffling the dataset to remove class consistency problems. Once, the data frame with shuffled, classes are balanced and consistent classes are formed, we then proceed to read the files as per the random shuffled pattern followed in the data frame. After reading the files as images of $128 \times 128 \times 3$ to give all the images dimensionality channels. We have come up with another data frame which consists of 50 thousand columns crossed over to nine thousand rows. Finally, the CSV file accounts for 4.67 GB storage space. Preprocessing stage also deals with normalization, we have used several normalizing techniques which are listed below the techniques used are:

1. After CSV Building
 - (a) Standard Scaler Normalization
 - (b) Normalization by statistical measures
 - (c) Min-Max normalization
 - (d) Decca-Norm.
2. Before CSV Building.
 - (a) dividing each image pixel by 255
 - (b) min-max scaling of pixels.

In order to reduce the computation, we have used principal dimensionality reduction, the following section describes the approach used (Fig. 5).

It can be seen clearly as per the PCA dimensionality approach that both the normal and pneumonia images aren’t much separable in two-dimensional spaces. However, it is also an experimental fact that we cannot plot more than three dimensions but on continuous experimentation and evaluation, we have been able to identify the most feasible and computationally efficient dimensional representation which spreads out over 8569 dimensions to find proper variance difference between the two classes.

After implementing the Machine Learning algorithm to test our dimensional reduction hypothesis. We have drawn up the confusion matrix, accuracy score, and the Classification Report (Fig. 6).

The accuracy seems reasonable for a machine learning algorithm. The classification report discloses the insights of class prediction as well as other parameters like precision, recall, F1 score, etc. (Fig. 7).

Fig. 5 Dimensionality reduction by principal component analysis

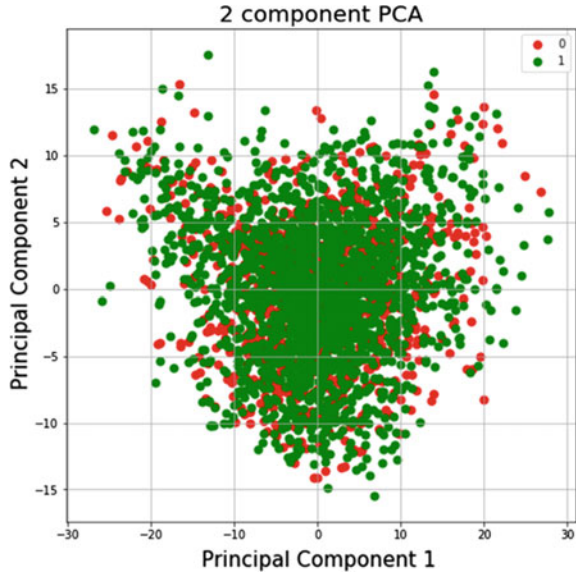


Fig. 6 Accuracy score

Accuracy Score 73.84969325153375

Fig. 7 Classification report

Classification Report					
	precision	recall	f1-score	support	
0.0	0.00	0.00	0.00	341	
1.0	0.74	1.00	0.85	963	
micro avg	0.74	0.74	0.74	1304	
macro avg	0.37	0.50	0.42	1304	
weighted avg	0.55	0.74	0.63	1304	

In the above classification report, we can see it conveying the class partiality which the algorithm has shown. The confusion matrix will tentatively show us a nominal prediction for the normal class (Fig. 8).

The above result indicates that even though the dimensionality reduction can be applied on blob segment analysis, the results are highly misleading and cannot be trusted for medical relics. Also, class learning has shown to be a severe problem. Also, it raises questions on the blob segmentation analysis that the blobs detected pertain to which type of blob.

Fig. 8 Distribution of results

```
Confusion Matrix [[ 0 341]
                  [ 0 963]]
```

1. Normal
2. Pneumonia
3. Any Other (Possibility of some other anomaly existing).

Clustering [15] is used to solve this conundrum as it is the best way to figure which blob segment belongs to which type. Clustering is a technique used to visualize similar points based on the boundary, statistical measures, and other algorithmic boundaries (Figs. 9 and 10).

Again, this seems relatable to the dimensionality PCA plots and also the classification results, the K-Nearest Neighbour clustering centroid showcases a relentless slope wherein all the points lie. So, it is understandable why dimensionality reduction at 8569 was of no use at all. Still, removing all the dimensionality out of the picture would be disastrous to the results.

On further investigation to improve the accuracy and other classification parameters. We have reached 12,288 dimensionalities and we can see our results as follows:- (Fig. 11).

Now, we have stuck to these very dimensions not only because of the results but also since further increasing the dimensions; leads to lagging in computational speed, decreasing accuracy and increasing time complexity.

Fig. 9 KNN elbow check

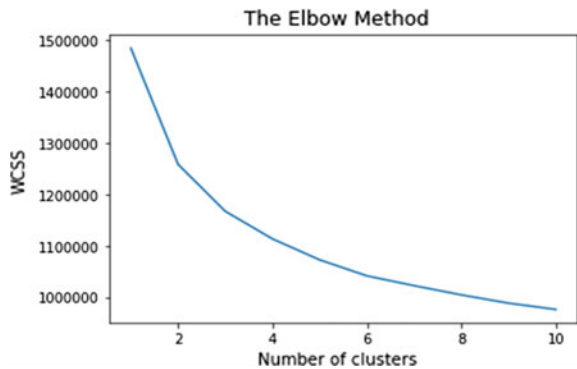


Fig. 10 KNN clustering

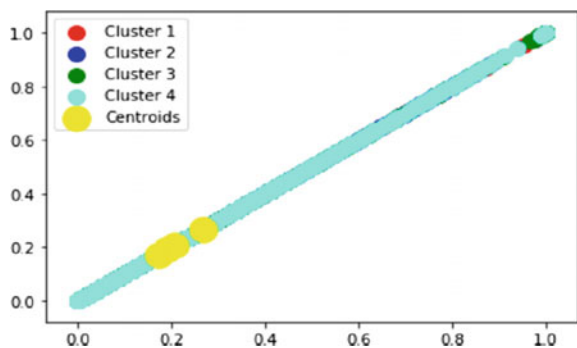


Fig. 11 Final results for clustering

```

Confusion Matrix [[300  48]
 [ 16 940]]
Accuracy Score 95.0920245398773
Classification Report
              precision    recall  f1-score   support

      0.0         0.95         0.86         0.90         348
      1.0         0.95         0.98         0.97         956

   micro avg         0.95         0.95         0.95        1304
   macro avg         0.95         0.92         0.94        1304
  weighted avg         0.95         0.95         0.95        1304
    
```

	Model Used	Overfir Quotient	Train Accuracy	Test Accuracy	F1 SCORE
0	Logistic Regression	0.011764	0.961917	0.950152	0.955943
1	SGDClassifier	0.019225	0.929412	0.910187	0.952740
2	SVM - Gamma --> scale	0.011267	0.965236	0.953969	0.969997
3	SVM - Gamma --> auto	0.048052	0.938397	0.890345	0.942822
4	NuSVM - Gamma --> scale	-0.014796	0.889331	0.904127	0.865739
5	LinearSVM - Gamma --> scale	0.012781	0.962170	0.949389	0.958955
6	KNeighborsClassifier	0.021240	0.926131	0.904891	0.928322
7	GaussianNB	-0.008926	0.849928	0.858854	0.857322
8	BernouliNB	-0.017923	0.710641	0.728564	0.742649
9	tree	0.035310	0.878838	0.843527	0.869632

Fig. 12 Genesis of machine learning algorithms

Thus, we have finally trained a number of models on 12,288 dimension parameters. Let’s have a look at this data frame to do further analysis (Fig. 12).

General algorithms have been used to generalize as well as garner better accuracies. We check with logistic regression and its results. It is intuitively brilliant to see a simplistic approach algorithm to showcase such good and generic results, the confusion matrix of the above algorithm is given as follows: (Figs. 13 and 14).

It showcases that the absence of very less false positives is watered down to 20 and also there is no presence of class partiality any further.

Next, Support Vector Machine (SVM) scaled by gamma. Following is its classification report.

This has given an even better result than that of logistic regression. We see an even lower false positive rate than the one in Logistic Regression (Figs. 15, 16).

The false-positive cases are lowered down to 16 and also there is no presence of class partiality further. This is even less than our previous best linear regression.

We would also like to turn to the possibility of gradient boosting algorithms providing better results down; which they have come to showcase in recent years. Their result simulations are listed below for further analysis and understanding.

```

Accuracy Score 95.62883435582822
Classification Report
              precision    recall  f1-score   support

   0.0         0.94      0.90      0.92       353
   1.0         0.96      0.98      0.97       951

 accuracy          0.96       1304
 macro avg         0.95       0.94       0.94       1304
 weighted avg      0.96       0.96       0.96       1304
    
```

Fig. 13 Logistic regression report

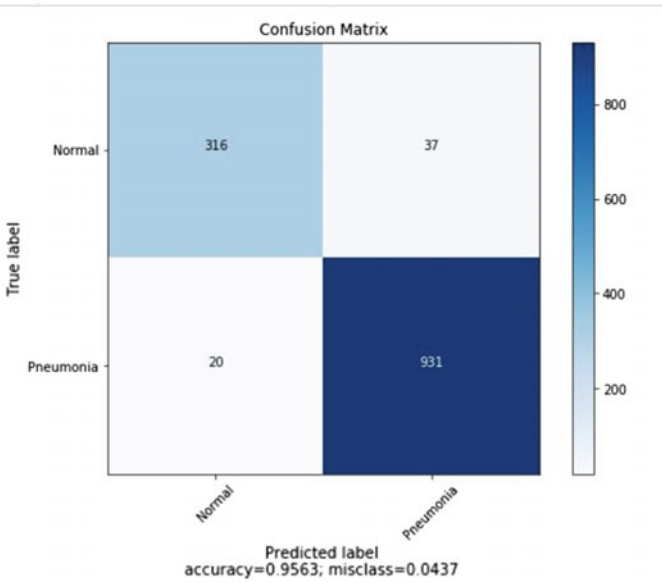


Fig. 14 Logistic regression distribution

```

Accuracy Score 97.00920245398773
Classification Report
              precision    recall  f1-score   support

   0.0         0.95      0.93      0.94       353
   1.0         0.98      0.98      0.98       951

 accuracy          0.97       1304
 macro avg         0.96       0.96       0.96       1304
 weighted avg      0.97       0.97       0.97       1304
    
```

Fig. 15 SVM report

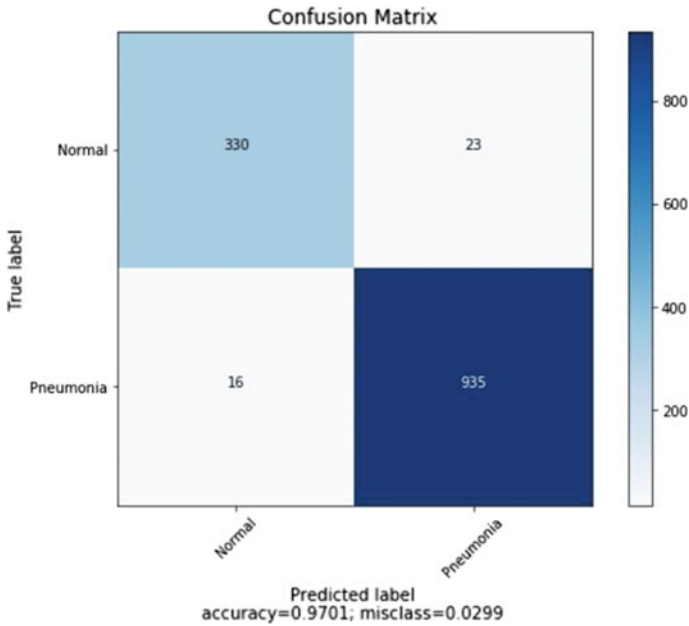


Fig. 16 SVM distribution

1. **Random Forest** (Figs. 17 and 18)
2. **Ada Boost** (Figs. 19 and 20)
3. Gradient Boosting (Figs. 21 and 22).

The gradient boosting [16] algorithms have also shown significant and generic results, the data structure and handling are also the same as described below.

It is important to realize that this complete training and testing has been handled and done in a time period of 29 h and 37 min. The system requirements are the same as described below in the deep learning and further algorithmic section.

```

-----Random Forest-----
Accuracy Score 93.55828220858896
Classification Report
      precision    recall  f1-score   support

     0.0         0.90     0.86     0.88         353
     1.0         0.95     0.97     0.96         951

 accuracy          0.94         1304
 macro avg         0.92         1304
 weighted avg     0.93         1304

```

Fig. 17 Random forest report

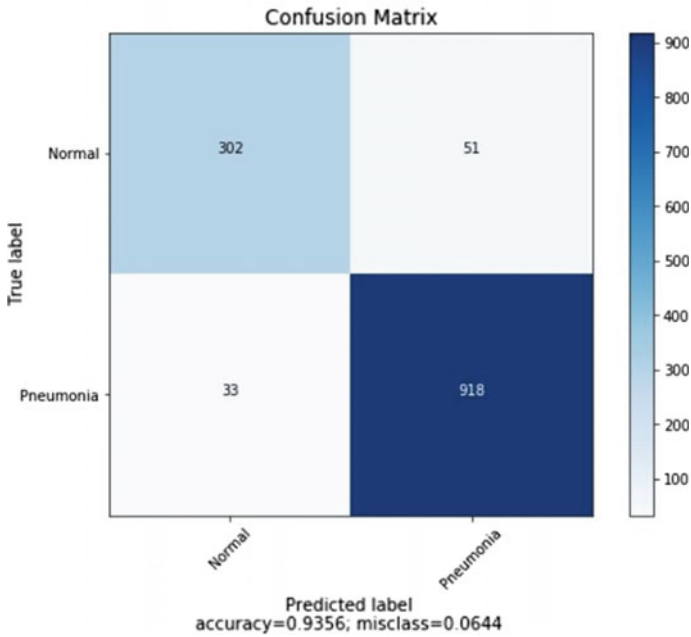


Fig. 18 Random forest distribution

```

-----Ada Boost-----
Accuracy Score 90.79754601226993
Classification Report
              precision    recall  f1-score   support

   0.0         0.87      0.77      0.82       353
   1.0         0.92      0.96      0.94       951

 accuracy          0.91       0.91       0.91      1304
 macro avg         0.90      0.87      0.88      1304
 weighted avg      0.91      0.91      0.91      1304

```

Fig. 19 Ada Boost report

6.2 Transfer Learning and Image Analysis

1. Mobile-net or Google Flow Model

We tried training the data through the Mobile-net/Google Flow model as shown below: (Fig. 23).

The biggest pitfall of this model is the low accuracy of the provided dataset. The model’s low accuracy helps us to lead to the conclusion that this would not be an appropriate solution to the problem presented.

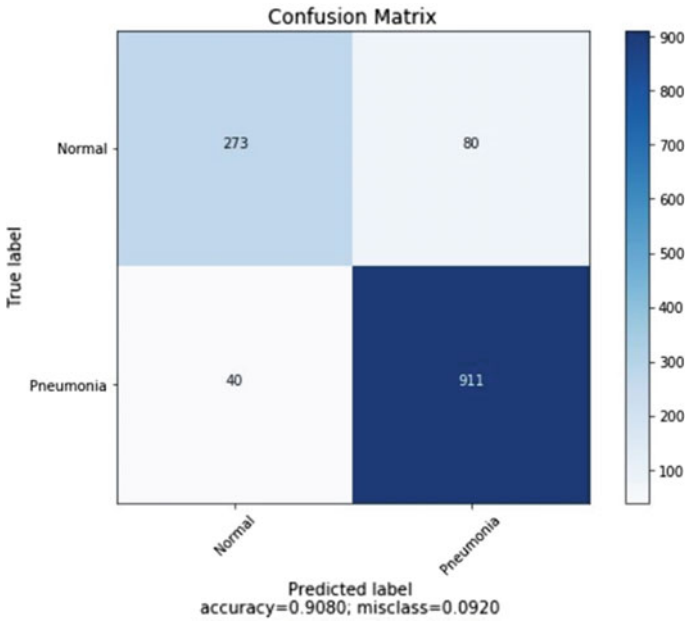


Fig. 20 Ada Boost distribution

```

-----Gradient Boosting-----
Accuracy Score 90.72085889570552
Classification Report

```

	precision	recall	f1-score	support
0.0	0.86	0.78	0.82	353
1.0	0.92	0.95	0.94	951
accuracy			0.91	1304
macro avg	0.89	0.87	0.88	1304
weighted avg	0.91	0.91	0.91	1304

Fig. 21 Gradient boost report

2. VGG-16

We also implemented the already available VGG-16 model and while it provided greater accuracy than the mobile-net model as can be seen in Fig. 24.

Even though the model has an increased accuracy score than the mobile-net model but a better accuracy score can be gained by a specifically designed neural network model [17] which can ensure less variation among the input layer and also the class imbalance is dealt in the model.

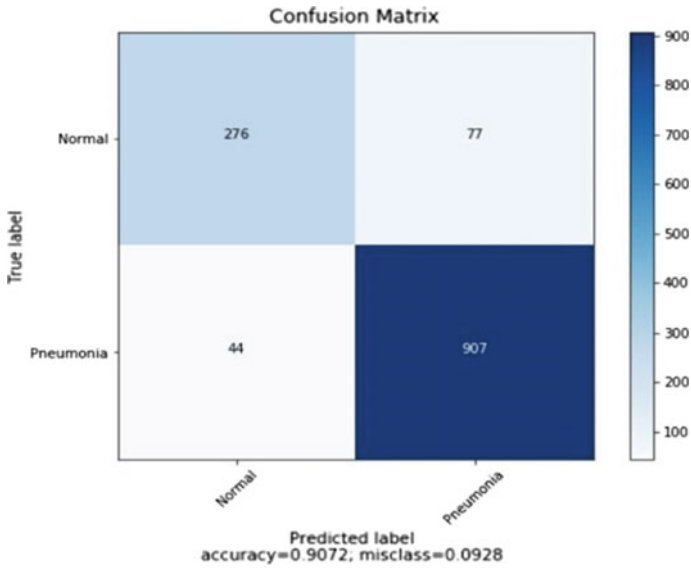


Fig. 22 Gradient boost distribution

```
Epoch 14/20
22045/22045 [=====] - 17s 779us/sample - loss: 0.6758 - accuracy: 0.5840
Epoch 15/20
22045/22045 [=====] - 17s 790us/sample - loss: 0.6724 - accuracy: 0.5850
Epoch 16/20
22045/22045 [=====] - 17s 773us/sample - loss: 0.6686 - accuracy: 0.5936
Epoch 17/20
22045/22045 [=====] - 17s 779us/sample - loss: 0.6657 - accuracy: 0.5996
Epoch 18/20
22045/22045 [=====] - 18s 808us/sample - loss: 0.6666 - accuracy: 0.5971
Epoch 19/20
22045/22045 [=====] - 17s 763us/sample - loss: 0.6647 - accuracy: 0.5997
Epoch 20/20
22045/22045 [=====] - 17s 791us/sample - loss: 0.6632 - accuracy: 0.6036

<tensorflow.python.keras.callbacks.History at 0x17c50af4da0>

1 model.evaluate(test_x, test_y)
5512/5512 [=====] - 2s 283us/sample - loss: 0.6706 - accuracy: 0.5813
[0.6705947292560417, 0.5812772]
```

Fig. 23 Mobile net

Fig. 24 VGG 16 report

	precision	recall	f1-score	support
0	0.65	0.45	0.53	2713
1	0.59	0.77	0.67	2799
3	0.00	0.00	0.00	0
micro avg	0.61	0.61	0.61	5512
macro avg	0.41	0.40	0.40	5512
weighted avg	0.62	0.61	0.60	5512

6.3 Proposed Neural Network and Model

Before the training of the model starts, we realize that the training dataset might cause class imbalance due to a lesser number of images being provided for ‘Normal’ Label.

To solve this, a **General Adversarial Network** is implemented.

For the first stage of this model, a Convolutional Layer which extracts the features with the help of kernels. These kernels are trained to detect specific features (or Trends) in the image. These kernels move over each part and detects whether the part has the specified resource or not.

If the required feature is present then the function returns a high valued real number otherwise it results in a low valued number being returned.

For a given Two-dimensional image I , and a small array, K of size $h \times w$ (kernel), the convoluted image ($I * K$), is calculated via superimposing the kernel/filter at the top of the image’s all possible forms, and recording the sum of the elementary products between the Image I and following kernel equation

$$(I * K)_{xy} = \sum_{i=1}^h \sum_{j=1}^w K_{ij} \cdot I_{x+i-1,y+j-1}$$

The output from this Convolutional Layer is transferred to the pooling layer which is responsible for reducing the spatial size of the feature map, preserving the resources detected into a smaller representation. Although, many alternatives are accessible for pooling. Max Pooling is the most. Max Pooling operates by locating the locations in the image that shows the strongest correlation with each resource (the maximum value) are preserved and these values combine to form a smaller space. As shown as Fig. 25.

After going through this process, we progress with **BatchNormalization** as explained in the concepts batch normalization helps with providing a normalized

Fig. 25 Max pooling



data which ranges from scale 0 to 1 for a data which earlier had a larger range. If in between normalizations are not done there might be one weight dominating others and affecting the results. The layer inputs also, as a result of the normalization, show less variation amongst them.

After Normalizing, the layer inputs which are scaled, are then processed through another Convolutional Neural Network and then through another Max Pooling which again reduces the spatial size of the feature map which is available. Then, we apply **Dropout** because it is a vital step for the enhancement of model which do not over fit, it randomly cancels out nodes in each iteration so that not anyone node is of greater importance than the others to the model. Also so that neither of single input features is heavily weighted, since under each backpropagation iteration the weights are decreased when dropped and failure of output occurs.

Then **Flatten** is used for converting the three-dimensional matrix to the one-dimensional matrix. This is done in order to make the data eligible to be fitted in artificial neural networks. However, it is of import to note that when training with more than one Convolutional.

Layers, flatten must be only used after the last convolutional layer. Once flatten is applied no more convolutional layers can be employed.

After these 2 dense layers are used. These dense layers directly refer to the Artificial Neural network schema which is 128 nodes in one layer. After the layers are added a Dropout is added and finally, another dense layer is added which has finally 2 nodes. Total Parameters are 331,330 while 331,266 are trainable parameters. For the compilation of the model, we use Sparse Categorical cross-entropy as the loss and 'Adam' as the optimizer.

The **cross-entropy** loss measures the performance of a classification model, with the output being a probability value ranging from 0 to 1.

Hyper parameter Information.

1. All dropouts are 25% that is one fourth so that overfitting can be removed.
2. For regularization, mostly Manhattan distance or L1 regularization is used.
3. The activation function is leaking rectilinear units and rectilinear units.
4. Adam Optimizer function is used to find the global minima and not to be stuck with the local minima.

Model Summary (Fig. 26).

The figure provided is a summary of the model we have implemented. It clearly shows the total number of trainable parameters.

7 Results

With the model compiled, now it is fitted on the dataset and accuracy of the model by using validation. Figure 27 shows the training of our model over 10 epochs.

As we can see the accuracy reaches 98.62%, this is because our model reduces the errors by utilizing the Batch Normalization and Drop Outs at the correct stages of the

```

Model: "sequential_19"

```

Layer (type)	Output Shape	Param #
conv2d_36 (Conv2D)	(None, 31, 31, 32)	896
max_pooling2d_22 (MaxPooling)	(None, 15, 15, 32)	0
batch_normalization_v2_29 (B	(None, 15, 15, 32)	128
conv2d_37 (Conv2D)	(None, 13, 13, 64)	18496
max_pooling2d_23 (MaxPooling)	(None, 6, 6, 64)	0
dropout_36 (Dropout)	(None, 6, 6, 64)	0
flatten_18 (Flatten)	(None, 2304)	0
dense_40 (Dense)	(None, 128)	295040
dense_41 (Dense)	(None, 128)	16512
dropout_37 (Dropout)	(None, 128)	0
dense_42 (Dense)	(None, 2)	258

```

=====
Total params: 331,330
Trainable params: 331,266
Non-trainable params: 64

```

Fig. 26 Model description

Fig. 27 Loss summary

```

loss: 0.5805 - accuracy: 0.7423 - val_loss: 0.7175
loss: 0.5547 - accuracy: 0.7433 - val_loss: 0.7571
loss: 0.5141 - accuracy: 0.7504 - val_loss: 0.7885
loss: 0.4518 - accuracy: 0.7872 - val_loss: 1.0575
loss: 0.3450 - accuracy: 0.8449 - val_loss: 1.0626
loss: 0.2253 - accuracy: 0.9141 - val_loss: 1.3736
loss: 0.1241 - accuracy: 0.9563 - val_loss: 2.0895
loss: 0.0780 - accuracy: 0.9768 - val_loss: 2.1105
loss: 0.0494 - accuracy: 0.9854 - val_loss: 2.5310
loss: 0.0593 - accuracy: 0.9862 - val_loss: 2.4531

```

network. Therefore, we have named this model as ‘Phoenix’, because as the myth of the phoenix suggests about it rising from the ashes after death, we believe our model arises from the depths of pre-processing errors, models and class imbalances. Our model overcomes all of these and hence provides the most accurate diagnoses while not being over fit.

The outputs of ‘Pneumonia’ and ‘Normal’ labels are encoded as 0 and 1 and as we can see the precision for both is quite high which validates our model even further and shows the exactness of our model.

Now, since we know the Phoenix Model has decent accuracy and is precise enough to be used in Medical Practice.

We now compare our activation functions for the layers in our network.

The following figure (Fig. 28) is the classification report of our model.

In Figs. 29 and 30 the plotted graphs showcase 3 activation functions, in comparison to each other plotted between the loss in training of train and validation sets of data respectively.

These plots depict clearly that the sigmoid activation function is of no use if the lower number of epochs is to be trained while the losses almost approach zero with the Tanh activation function which again points us to the conceptual error that losses cannot be almost zero since in real-time.

Similar conclusions as above can be made from Figs. 31 and 32 of activation functions on accuracy versus epoch graphs on both training and validation sets.

Next, we move on and find the best optimizer for the dataset. We have used ‘Adam’ optimizer in our model.

Again, we plot for training and validation sets loss against the epochs trained. In Fig. 33 the respective graphs are present for four optimizers.

These plots show that ‘Adam’ as an optimizer has the lowest losses amongst all the other optimizers which in turn shows that ‘Adam’ is the best choice for the optimizer.

Fig. 28 Report

	precision	recall	f1-score	support
0	1.00	0.87	0.93	670
1	0.90	1.00	0.95	766
accuracy			0.94	1436
macro avg	0.95	0.93	0.94	1436
weighted avg	0.94	0.94	0.94	1436

Fig. 29 Activation function training loss

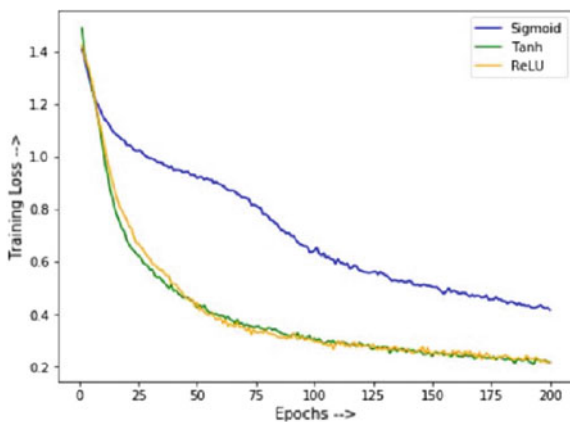


Fig. 30 Activation function test loss

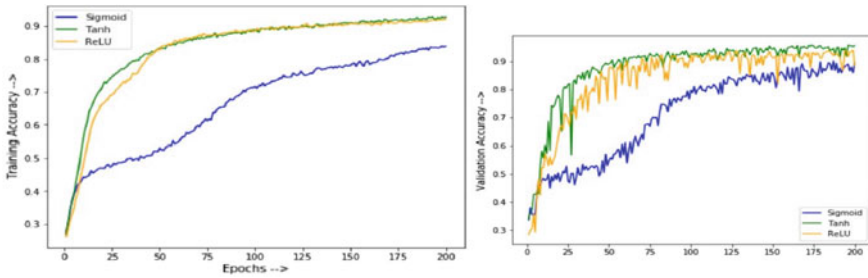
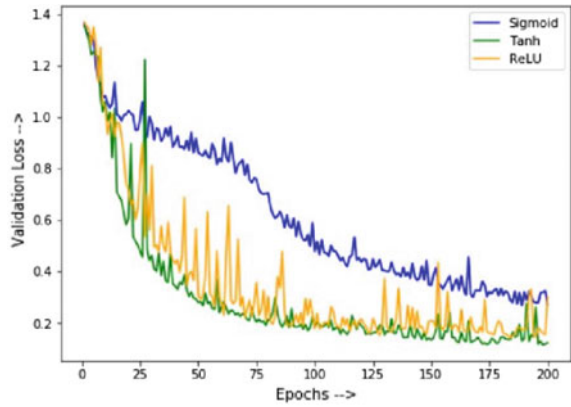


Fig. 31 Activation function accuracy

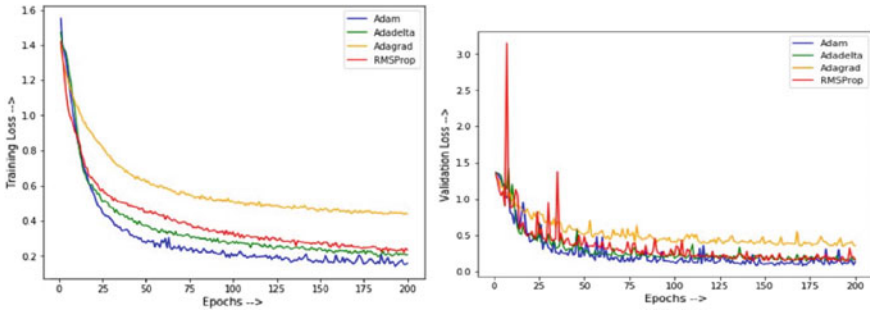


Fig. 32 Optimizer loss

Further, we plot the difference between our train and test accuracy and we showcase that the Phoenix model has high accuracy and comparatively less loss. This is depicted in the following Figs. 34 and 35.

As is clear from these graphs the model's accuracy is high and loss is relatively low. It is due to various factors which we have included in our model such as.

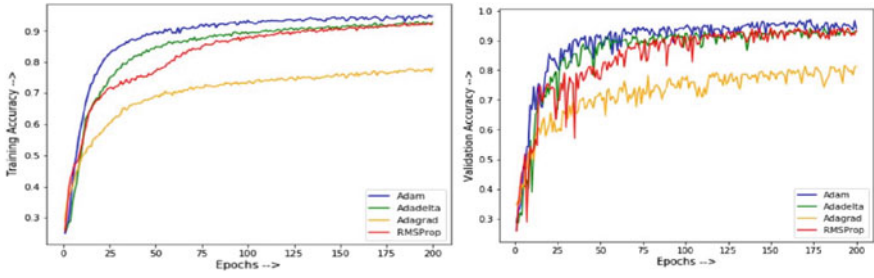


Fig. 33 Optimizer accuracy

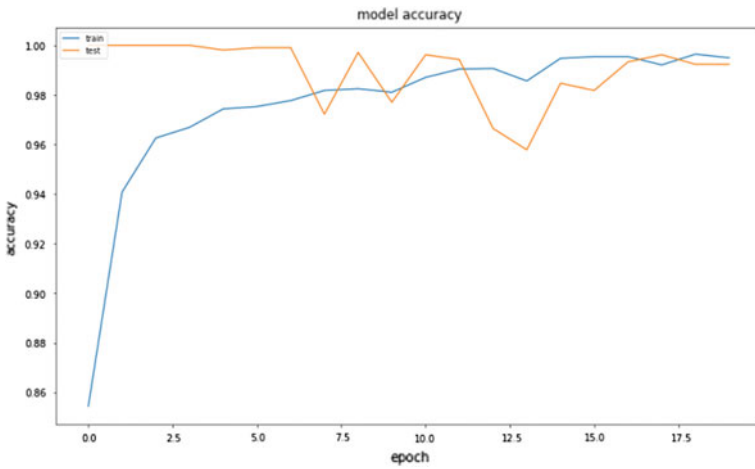


Fig. 34 Model accuracy

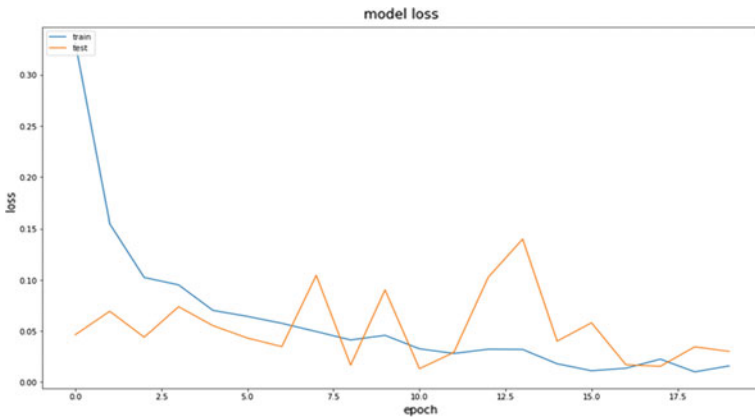


Fig. 35 Model loss

1. GANs: which help us address the problems of class imbalance due to the dataset having it inherently
2. Batch Normalization so that layer inputs have a varied distribution.
3. Max Pooling after each Convolutional Neural Network layer.
4. Usage of ‘relu’ as activation function and ‘Adam’ as an optimizer.

Confusion Matrix.

The confusion matrix is an array that contains correct and incorrect predictions of the algorithm and the actual situation.

- True Positive: Number of people who actually have pneumonia according to the algorithm.
- False Negative: Number of people who are actually with pneumonia but categorized as healthy according to the algorithm.
- False Positive: Number of people who are actually healthy, but categorized as pneumonia, according to the algorithm.
- True Negative: Number of people who are really healthy and categorized as healthy according to the algorithm

Figure 36 depicts the confusion matrix for our model Phoenix which shows that the True Negatives and True Positives are classified almost correctly according to the provided.

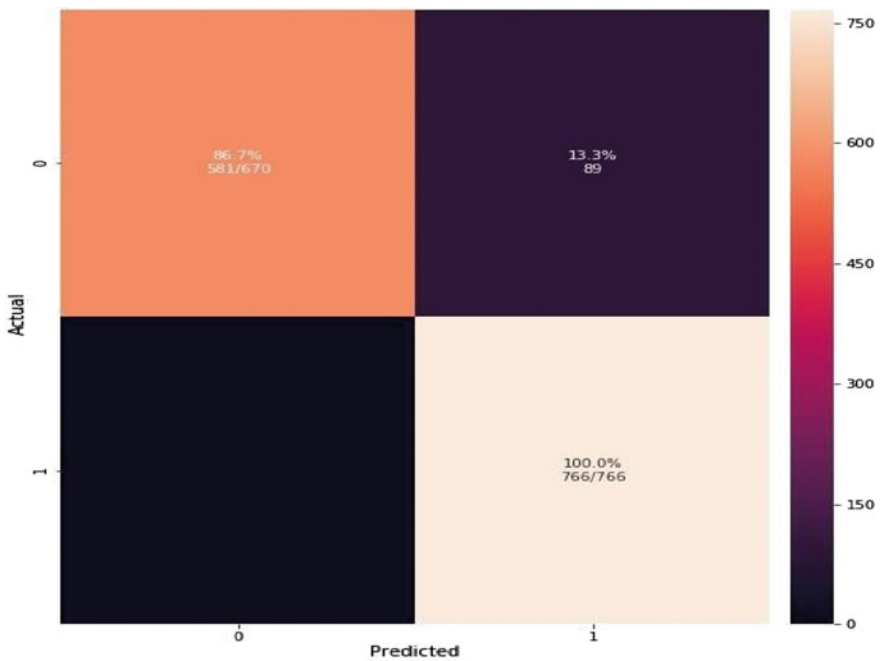


Fig. 36 Model report distribution

8 Conclusion

In this paper ‘Phoenix’ Model has been proposed for diagnosing Pneumonia through Chest X-rays and computer vision with a neural network model being trained. We have achieved 98.6% accuracy on the (Kermany) dataset on which earlier researches have proposed a maximum of average accuracy of 95.30% [8] and 92.8% (Kermany et al.). We believe that our model solves the predicament we found ourselves in earlier while implementing Mobile-net and VGG-16 quite efficiently and has quite a significant rise in terms of accuracy and precision. In comparison to the machine learning algorithms which tested SVC 97.001% and Logistic regression 95.643%, our model (Phoenix) is better suited for the purposes of the medical industry.

Both the models showcase false positive error of amount 16 and 20 per test sample input cases respectively. As described earlier our version of evaluating the post matrice results was to decrease the false positive so that medical robustness can be achieved. Further strengthening that note we are able to deduce about 0 true negatives using our phoenix model on the given input test sequence. Finally with evidence from both machine learning and deep learning relics of data analytics. Further enhancement and decreasing the false positives and increase of accuracy as well as other parameters of evaluation compared to the earlier models.

References

1. H. Ketout, MVN_CNN and UBN_CNN for endocardial edge detection (2011)
2. S. Khobragade, A. Tiwari, C.Y. Patil, V. Narke, Automatic detection of major lung diseases using chest radiographs and classification by feed forward artificial neural network, in *2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, (2016), pp. 1–5
3. A. Khan, A. Khalil, H. Hajjdiab, Mobile microscopic device to detect parasitical cell related diseases using machine learning
4. S. Mandal, V.E. Balas, R.N. Shaw, A. Ghosh, Prediction analysis of idiopathic pulmonary fibrosis progression from OSIC dataset, in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, 2–4 Oct. 2020, pp. 861–865. <https://doi.org/10.1109/GUCON48875.2020.9231239>
5. R. Barrientos, A. Roman-Gonzalez, F. Barrientos, L. Solis, M. Correa, M. Pajuelo, C. Anticon, R. Lavarello, B. Castaneda, R. Oberhelman, W. Checkley, R. H. Gilman, M. Zimic, Automatic detection of pneumonia analyzing ultrasound digital images, in *2016 IEEE-36th Central American and Panama Convention (CONCAPAN XXXVI)* (2016), pp. 1–4
6. M.B. Rodrigues, R.V.M.D. Nobrega, S.S.A. Alves, P.P.R. Filho, J.B.F. Duarte, A.K. Sangaiah, V.H.C.D. Albuquerque, Health of things algorithms for malignancy level classification of lung nodules. *IEEE Access* **6**, 18592–18601 (2018)
7. H.Y. Jia, J.C.S. Nowa, H. Long Y.D. Liu, Soil fertility with bayesian network of transfer learning (2010)
8. A. Saraiva, N. Ferreira, L. Lopes de Sousa, N. Carvalho da Costa, J. Sousa, D. Santos, A. Valente, S. Soares, Classification of images of childhood pneumonia using convolutional neural networks, 112–119 (2019). <https://doi.org/10.5220/0007404301120119>
9. S. Kido, Detection of lung abnormality by use of Convolutional Neural Networks and Region CNN’s (2018)

10. T. Takano, H. Takase, H. Kawanaka, S. Tsuruoka, Transfer learning for reinforcement learning in the same transition model quick approach and preferential exploration, in *2011 12. 12. 10th International conference on Machine Learning and Application and Workshops*, vol. 1 (2011)
11. M. Kumar, M. Shenbagaraman, R.N. Shaw, A. Ghosh, Predictive data analysis for energy management of a smart factory leading to sustainability, in M. Favorskaya, S. Mekhilef, R. Pandey, N. Singh (eds.) *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol. 661 (Springer, Singapore 2021). https://doi.org/10.1007/978-981-15-4692-1_58
12. Predictive Analysis. <https://in.mathworks.com/discovery/predictive-analytics.html>
13. Pneumonia Deaths under the age of five UNICEF. https://www.unicef.org/health/index_91917.html
14. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, General Adversarial Nets , in *Conference on Advances in Neural Information Processing Systems* (2014)
15. S. Yue, S. Ozawa, A sequential Multi task learning network with metric based knowledge transfer (2012)
16. S. Mandal, S. Biswas, V.E. Balas, R.N. Shaw, A. Ghosh, Motion prediction for autonomous vehicles from lyft dataset using deep learning, in *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)* 30–31 Oct. 2020, pp. 768–773. <https://doi.org/10.1109/ICCCA49541.2020.9250790>
17. Shu, G.S., W. Liu, X. Zheng, J.L. IF-CNN, Image Aware Inference Ling Shao, Transfer Learning for Visual Categorization: A Survey (2015)
18. Y.M. Hirimuthugoda, Artificial intelligence based approach for determination of haematologic diseases (2009)
19. K. Sethi, V. Parmar, M. Suri, Low power hardware based deep learning diagnostics support case study (2018)